Plant Cell Advance Publication. Published on July 21, 2017, doi:10.1105/tpc.17.00354

RESEARCH ARTICLE

Differentially regulated orthologs in sorghum and the subgenomes of maize

Yang Zhang[1,2], Daniel W. Ngu[1,2], Daniel Carvalho[1,2], Zhikai Liang[1,2], Yumou Qiu[3], Rebecca L. Roston[1,4], and James C. Schnable[*,1,2]

[1] Center for Plant Science Innovation, University of Nebraska-Lincoln, Lincoln, Nebraska, USA

[2] Department of Agronomy and Horticulture, University of Nebraska-Lincoln, Lincoln, Nebraska, USA
[3] Department of Statistics, University of Nebraska-Lincoln, Lincoln, Nebraska, USA
[4] Department of Biochemistry, University of Nebraska-Lincoln, Lincoln, Nebraska, USA
* Corresponding Author: schnable@unl.edu

Short title: Gene regulatory changes between maize and sorghum

One-sentence summary: The response of the same genes to cold stress often varies between maize and sorghum, but the set of genes with conserved patterns of regulation show greater evidence of functional constraint.

The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantcell.org) is: James C. Schnable (schnable@unl.edu).

ABSTRACT

Identifying interspecies changes in gene regulation, one of the two primary sources of phenotypic variation, is challenging on a genome-wide scale. The use of paired time course data on cold-responsive gene expression in maize (*Zea mays*) and sorghum (*Sorghum bicolor*) allowed us to identify differentially regulated orthologs. While the majority of cold-responsive transcriptional regulation of conserved gene pairs is species specific, the initial transcriptional responses to cold appear to be more conserved than later responses. In maize, the promoters of genes with conserved transcriptional responses to cold tend to contain more micrococcal nuclease hypersensitive sites in their promoters, a proxy for open chromatin. Genes with conserved patterns of transcriptional regulation between the two species show lower ratios of nonsynonymous to synonymous substitutions. Genes involved in lipid metabolism, known to be involved in cold acclimation, tended to show consistent regulation in both species. Genes with species-specific cold responses did not cluster in particular pathways nor were they enriched in particular functional categories. We propose that cold-responsive transcriptional regulation in individual species may not be a reliable marker for function, while a core set of genes involved in perceiving and responding to cold stress are subject to functionally constrained cold-responsive regulation across the grass tribe Andropogoneae.

1

INTRODUCTION

The grasses are a clade of more than 10,000 species, which exhibit conserved morphology and genome architecture (Bennetzen and Freeling, 1993). Grasses have adapted to grow in a wide range of climates and ecologies across the globe, with 20% of total land area covered by ecosystems dominated by grasses (Shantz, 1954). As a result, the range of tolerance to abiotic stresses present in the grass family (Poaceae) far exceeds that present within any single grass species. However, to date, studies attempting to identify determinants of abiotic stress tolerance at a genetic or genomic level have predominantly focused on individual species (Chopra et al., 2017; Priest et al., 2014; Revilla et al., 2016; Tiwari et al., 2016; Waters et al., 2017). The majority of genetic changes with phenotypic effects can be broadly classified into two categories: those that alter protein-coding sequence and those that alter the regulation of gene expression.

DNA sequence changes that alter protein-coding sequences can be identified in a straightforward fashion. The probability that a given polymorphism in a protein-coding sequence will have a phenotypic effect can also often be estimated. At a basic level, this involves classification as synonymous, missense and nonsense mutations. Information on the overall level of evolutionary conservation for a given amino acid residue can also be used to increase the accuracy of these predictions (Cooper et al., 2005; Ng and Henikoff, 2001; Reva et al., 2011). Cross species comparisons of the protein-coding sequences from genes co-opted into new functional roles in C4 photosynthesis have been able to identify protein changes linked to changes in function at a resolution of individual amino acid residues (Christin et al., 2007).

Identifying changes in gene regulation across related species is more challenging, and the associated methods are far less advanced. For extremely close relatives, such as *Arabidopsis thaliana* and *Arabidopsis arenosa*, RNA-Seq reads from both species can be mapped to a common reference genome (Burkart-Waco et al., 2015). For species with greater levels of sequence divergence in transcribed regions, this approach becomes impractical. Recent work in Sophophora (formerly Drosophila) described some of the many challenges present in comparing changes in baseline expression levels across closely related species with independently sequenced and assembled reference genomes (Torres-Oliva et al., 2016). However, this approach is limited to identifying changes in baselines expression in the same treatment rather than

examining patterns of regulation across multiple treatments. Within the grasses several research groups have employed clustering-based methods to identify genes with conserved patterns of regulation during either reproductive or photosynthetic development (Davidson et al., 2012; Wang et al., 2014). Among other results, one of these studies concluded that orthologous genes conserved at syntenic locations are more likely to share correlated expression patterns across multiple species than genes classified as orthologs based on phylogenetic analysis but located at non-syntenic locations (Davidson et al., 2012). Clustering-based methods can identify genes with conserved patterns of regulation across multiple species, but they have high false positive rates when used to identify genes with changes in regulatory pattern.

In even closely related species, the baseline expression levels of orthologous genes can diverge significantly (Hollister and Gaut, 2009; Hollister et al., 2011). Testing for conserved or divergent patterns of regulation across different genotypes or different species when baseline expression levels have diverged creates a statistical challenge. Modeling of multiple environmental or genotype level effects can be combined either additively or multiplicatively. The model selected will determine which set of genes will be classified as differentially regulated between species. While few attempts have been made to identify differential patterns of gene regulation across species, attempts to do so between subspecies or diverse accessions have largely used either only a multiplicative model (Lovell et al., 2016), an additive model, or additive and multiplicative models separately (Waters et al., 2017) but have not made comparisons between the suitability of the two models.

Here, we sought to develop effective methods for comparing gene regulatory patterns between syntenic orthologous genes in closely related species. For initial cross species comparisons, data on changes in the transcriptional responses to cold stress in maize and sorghum were employed. Cold was selected as a stress that could be delivered in a consistent fashion and time frame. Maize and sorghum were selected based on their close evolutionary relationship (Swigoňová et al., 2004), high quality sequenced genomes (Paterson et al., 2009; Schnable et al., 2009), and common susceptibility to cold stress (Chinnusamy et al., 2007; Hetherington et al., 1989; Wendorf et al., 1992). In addition, maize is a mesotetraploid species that experienced a whole-

genome duplication approximately 12 million years ago after its divergence from sorghum (Swigoňová et al., 2004), producing two functionally distinct maize subgenomes, maize1 and maize2 (Schnable et al., 2011). Approximately 3,000-5,000 pairs of genes are retained on both maize subgenomes (Schnable et al., 2011, 2012, 2009). Unlike other types of gene duplication, whole-genome duplicates initially retain almost all the same associated conserved regulatory sequences (Freeling et al., 2012). Comparing the expression patterns of duplicated genes exposed to the same trans-regulatory factors provides a bridge to comparing the expression patterns of orthologous genes in closely related species with similar phenotypes. These two systems provide a useful platform for developing and testing approaches to comparative gene regulatory analysis. However, one goal of cross species comparisons of transcriptional regulation must ultimately be to link changes in regulation to changes in phenotype, which in the case of low temperature stress will require conducting comparisons between species with differing, rather than similar, tolerance to cold.

RESULTS

A set of 15,231 syntenic orthologous gene pairs conserved between the maize1 subgenome and sorghum and 9,553 syntenic gene pairs conserved between the maize2 subgenome were employed in this study (Figure 1A). The sequence identity in coding regions of syntenic genes between sorghum and either maize subgenome or between maize subgenomes is approximately 90% (Supplemental Figure 1), which is a level of divergence that makes alignment to a common reference sequence impractical. We conducted parallel expression analyses of the set of syntenic orthologous gene pairs conserved between the maize1 subgenome and sorghum and the smaller set of syntenic gene pairs conserved between the maize2 subgenome and sorghum.

Syntenic orthologs exhibited reasonably well-correlated patterns of absolute gene expression levels between sorghum and either subgenome of maize based on expression data generated from whole seedlings under control conditions (Spearman's rho = 0.79-0.84, Pearson r = 0.67-0.85, Kendall rank correlation 0.67-0.63, Figure 1B). This observation is consistent with previous reports about the analysis of expression across reproductive tissues in three grass species (Davidson et al., 2012). However, it should be noted that these correlations were significantly lower than those observed between biological replicates (see Methods for a detailed explanation

4

of what constituted a biological replicate in this study) of the same species (Spearman's rho = 0.88-0.98, Pearson r = 0.89-0.99, Kendall rank correlation 0.78-0.91), and many individual genes have large divergence in baseline expression levels between the two species, creating divergence between the predictions of additive and multiplicative statistical models of gene regulation, as described above.

We visually confirmed the lethal effect of prolonged cold stress on maize and sorghum (Ercoli et al., 2004; Hetherington et al., 1989; Olsen et al., 1993; Sánchez et al., 2014; Shaykewich, 1995) following prolonged cold treatment (Figure 2A-C, Supplemental Figure 2, See Methods). We employed measurements of impairment of CO_2 assimilation rates after recovery from a controlled length cold stress to provide more quantitative measures of cold stress and to assess the suitability of the level of cold stress employed to distinguish differing degrees of cold stress sensitivity or cold stress tolerance among maize, sorghum, and several related panicoid grass species. Data were generated from a total of six panicoid grasses, including the relatively cold tolerant paspalum (*Paspalum vaginatum*) and the extremely cold sensitive proso millet (*Panicum miliaceum)* (Figure 2D). After one day of cold stress, the species could be broadly classified as either cold stress insensitive or cold stress sensitive, with both maize and sorghum in the cold stress sensitive category. A longer period of cold stress (3 days) revealed greater impairment of CO_2 assimilation rates in sorghum than in maize, consistent with previous reports on the relative cold sensitivity of these two species (Chinnusamy et al., 2007; Chopra et al., 2017; Fiedler et al., 2016; Hetherington et al., 1989; Wendorf et al., 1992), and separated the six species into three broad categories of cold tolerant, moderately cold sensitive and extremely cold sensitive. Based on these data, we selected one day of cold stress, when maize and sorghum still exhibit comparable levels of CO_2 assimilation impairment (Figure 2D), for downstream expression analysis.

Conventional differentially expressed gene analysis

We identified differentially expressed genes in each species by comparing gene expression data in control seedlings to those subjected to one day of cold stress (Supplemental Data Set 1). Among maize1/sorghum syntenic gene pairs, 1,686 (11.1%, 1,686 out of 15,231) and 2,343 (15.4%, 2,343 out of 15,231) genes were classified as differentially expressed genes (DEGs),

respectively (Figure 3A, see Methods). For maize2/sorghum syntenic gene pairs, these values were 968 (10.1%, 968 out of 9,553) and 1,446 (15.1%, 1,446 out of 9,553) genes, respectively. Only 836 (5.5%, 836 out of 15,231) of maize1/sorghum syntenic genes were classified as showing differential regulation in response to cold in both species (Figure 3A). In addition, there were 29 and 16 genes pairs in the maize1/sorghum and maize2/sorghum gene pairs, respectively, where both genes were classified as differentially expressed but in opposite directions (Figure 3B). The 836 observed syntenic gene pairs is approximately 3.2 times higher than the 259 genes pairs that should have been identified if cold-responsive gene regulation were not correlated between the two species (see legend of Figure 3 for a detailed breakdown of how this value was calculated). With these two values, the maximum number of genes responding to cold in the same fashion as a result of common descent from an ancestrally cold-responsive gene in the common ancestor of maize and sorghum can be calculated using the formula ((observed number of shared DEGs) - (expected number of shared DEGs))/(observed number of shared DEGs). In this case, a maximum of approximately two thirds (69.0%, 577 out of 836) of genes identified as responding to cold in both species are likely to do so as a result of common descent. However, this may in fact be an overestimate if some of the same changes in cold-responsive gene regulation have been selected for in parallel in both lineages. Extending this calculation to the set of gene pairs that responded transcriptionally to cold in either maize or sorghum or both, only 18.1% (577 out of 3,193) of gene pairs responding to cold in either species are likely to have retained a conserved pattern of cold-responsive gene expression since the divergence of maize and sorghum from a common ancestor 12 million years ago (Swigoňová et al., 2004).

One potential explanation for this observation is that low statistical power to detect differentially expressed genes may create a false impression that differential expression is not conserved between related species. Prior estimates from real biological data in yeast (*Saccharomyces cerevisiae)* suggest that, given the number of replicates and minimum cut-off for differential expression employed here, the power of DESeq2 to identify differentially expressed genes should be between 0.65 and 0.90 (Schurch et al., 2016). In addition, a simulation study using observed expression values and variances in the maize dataset generated here indicated that the power to detect differential gene expression ranged from 0.63 for genes with a change in expression exactly at the minimum cut off to 0.961 for genes with larger changes in expression

6

value (Supplemental Data Set 2). The expected proportion of genes classified as differentially expressed in either species that are classified as differentially expressed in both species is given by the formula $power^2/1-(1-power)^2$. Given the worst-case assumption (power = 0.628), this value would be 46% if gene regulation were perfectly conserved between maize and sorghum, which is higher than the observed value of 25%.

Results for maize2/sorghum gene pairs were largely comparable. However, the proportion of genes classified as not differentially expressed in either species was greater for maize2/sorghum gene pairs (Figure 3A), likely because maize2 genes tend to have lower overall levels of expression (Schnable et al., 2011). In total, 766 nonsyntenic maize genes were classified as differentially expressed in response to cold (2.0% of all nonsyntenic genes in maize, 766 out of 38,664), while 1,333 (9.1%, 1,333 out of 14,683) of nonsyntenic genes in sorghum were classified as differentially expressed in response to cold. The absolute numbers of differentially expressed nonsyntenic genes are more similar to each other than the proportions, as the current set of maize gene model annotations includes many lower confidence genes - which are generally nonsyntenic and often show little or no detectable expression (Schnable, 2015) - than the current set of sorghum gene model annotations.

Maize and sorghum share a close relationship (Swigoňová et al., 2004), and both originated from tropical latitudes (De Wet, 1978; van Heerwaarden et al., 2011). The two species even have a high degree of promoter conservation in abiotic stress-responsive genes (Freeling et al., 2007). Therefore, the apparent low degree of conservation in cold stress-responsive regulation is unexpected. However, this result is also consistent with studies that have found significant divergence in abiotic stress responses between different haplotypes in maize (Waters et al., 2017).

One potential explanation is that the same cold stress pathways are being induced in maize and sorghum, but these pathways are induced more rapidly in one crop than the other when exposed to equivalent cold stresses. To test this hypothesis, we used data from a more detailed time course to compare the expression levels between matched pairs of cold stressed and control plants of each species at six time points distributed over 24 hours (See Methods and Supplemental Data Set 1). The number of gene pairs classified as differentially expressed at

different time points ranged from 60 to 2,199 for maize1/sorghum gene pairs and 29 to 1,235 for maize2/sorghum gene pairs. Comparing the number of genes identified as differentially expressed in each of all 36 possible pairwise combinations of time points between the two species showed that the greatest proportion of shared differentially expressed gene pairs was identified when identical time points were compared between the two species and that the overall number of shared differentially expressed gene pairs increases at later time points (Figure 4A). Overall, genes tended to remain in the same categories, with a general trend towards more DE0 genes moving into all three cold-responsive expression categories as the length of cold stress increased (Figure 4A). Because the proportion of all genes classified as differentially expressed increases at later time points, the expected number of gene pairs classified at DE2 under the null model described above also increases. Therefore, considering only the absolute number of gene pairs classified as DEGs in both species (DE2) at each time point can be misleading. After controlling for the expected number of DE2 genes, early time points show significantly higher proportions of true positives than later time points (Figure 4B).

Differentially regulated ortholog analysis

Another potential explanation for the finding that relatively few shared differentially expressed genes were identified between maize and sorghum is that differential gene expression analysis may not be testing the correct null hypothesis for between-species comparisons (Paschold et al., 2014). The null hypothesis of conventional DEG analysis is that the expression values observed for a given gene under control and stress conditions are drawn from the same underlying distribution. This approach is perfectly suitable for single-species analysis. In a two-species analyses, such as those conducted above, a DEG approach divides gene pairs into three categories: genes pairs classified as differentially expressed in neither species (DE0), in one species but not the other (DE1), and in both species (DE2, Figure 3A).

As shown in Figure 5A, in principle, each of those three categories (DE0, DE1, and DE2) can include gene pairs without significant differences in the pattern of regulation between species (comparably regulated orthologs or CROs), as well as gene pairs that do show significant differences in regulation between the two species (differentially regulated orthologs or DROs). All six theoretical cases from Figure 5A were observed in the RNA-Seq expression dataset

8

generated above (Supplemental Figure 3A). DROs and CROs were both observed in all the DEG groups (Supplemental Figure 3B). Distinguishing between DROs and CROs requires testing a different null hypothesis: that the change in expression for a given gene between two treatments is equivalent to the change in expression for an ortholog of that same gene, in a different species, across the same two treatments. Another way of describing this same experimental approach is testing for a statistically significant treatment by species interaction effect. Several existing statistical packages incorporate the ability to test for significant interactions between different treatments (Love et al., 2014; Ritchie et al., 2015; Robinson et al., 2010) by including species as an effect in the model. However, comparing across species under different conditions, including testing for interaction effects to cross species comparisons, requires us to define an accurate model for what the same change in gene regulation looks like starting from different baseline levels of expression. Testing this null hypothesis across species in turn requires us to define an accurate model of what the same pattern of gene expression looks like when starting from different baseline levels of expression.

For an orthologous gene pair where gene copies are expressed at different baseline levels in two species, two different models can be used to compare a change in expression between treatment and control conditions: additive and multiplicative (Figure 5B). When expression under control conditions is equivalent between the two species, these models yield the same predicted expression under stressed conditions. However, when control condition expression is different between the two species, the models produce different expected expression values under stress conditions. Using simulated data based on additive and multiplicative models, an ANOVA-based test classified genes with different baseline expression levels but the same pattern of expression (as simulated by a multiplicative model) as significantly differentially regulated between species, while the generalized linear model-based DESeq2 classified genes with different baseline expression levels but the same pattern of expression (as simulated by a additive model) as significantly differentially regulated between species (Supplemental Data Set 3).

To test which of these models is a better representation of how cold-responsive gene regulation actually operates, we utilized a set of 5,257 gene pairs retained from the maize whole genome duplication (WGD) (Schnable et al., 2011). The maize WGD created two copies of each gene in

the genome, each associated with the same chromatin environments and regulatory sequences. RNA-Seq-based measurements of expression for duplicate genes can be unreliable when gene copies are similar enough that reads cannot be unambiguously mapped to individual copies. Maize WGD-derived duplicate gene pairs show approximately 93% sequence similarity in exon regions (Supplemental Figure 1). This is equivalent to 4.5 mismatches per 50 bp sequence read, significantly reducing the risk of ambiguous or incorrect read mapping. The expression level of each gene copy in a WGD gene pairs in the maize genome in the same samples results from the exact same trans-factors acting in the exact same tissue and cell types. Therefore, divergence in the regulation of these genes should start out with the same cis-regulatory sequence prior to their divergence from their most recent common ancestor (whether at the time of WDG for autopolyploids or at the time of speciation prior to WGD for allopolyploids) (Freeling et al., 2012).

To test the additive and multiplicative null models, we used the expression pattern of one maize gene copy between control and cold stress conditions to predict the expression pattern of the other maize gene copy using each null model from Figure 5B. We conducted the analysis in parallel at each of the six time points in maize using maize1/maize2 gene pairs where at least one copy was identified as differentially expressed at that time point. Gene pairs were omitted from the analysis if the predictions of both models were more similar to each other than either was to the observed value.

The multiplicative model was more accurate at predicting cold-responsive expression patterns between maize WGD duplicates than the additive model at all time points (p=0.004-2.4*10-15), paired two tailed t-test) (Supplemental Data Set 4). Requiring the difference between the predictions of the two models to be at least twice as large as the difference between the better model and the observed expression pattern produced similar results (Figure 5C, Supplemental Data Set 4). The set of genes where the additive model produced better predictions was examined for differences in expression, selection (Ka/Ks ratio), (Supplemental Figure 4) or GO annotation No significant markers for which genes could be best predicted with which model were identified. Therefore, going forward, we employed the multiplicative model for conserved gene regulation across species, as implemented in DESeq2's test for multiple factors (Love et al.,

279 2014) (see Methods).
280
281 Figure 6A shows the proportion of gene pairs classified as DROs among all gene pairs in the
282 DE0, DE1, and DE2 groups at each of the six time points. Comparing the same time points for
283 maize and sorghum identifies fewer differentially regulated orthologs than comparisons between
284 non-equivalent time points in the two species. Fewer differentially regulated orthologs were
285 identified at earlier cold treatment time points than at later time points. This is consistent with the
286 results of DEG analysis described above, which suggested early cold stress responses were more
287 conserved across sorghum and maize than later cold stress responses.
288
289 **Functional differences between genes with conserved or lineage-specific regulatory**
290 **patterns**
291 Genes classified as responding to cold stress in both species (DE2) tended to have significantly
292 lower ratios of nonsynonymous nucleotide changes to synonymous nucleotide changes (Ka/Ks
293 ratio) than genes that responded to cold stress in only one species or in neither species. This
294 suggests genes with conserved patterns of cold-responsive regulation experience stronger
295 purifying selection than genes with lineage-specific patterns of cold-responsive regulation
296 (Figure 6B-C). GO (Gene Ontology) enrichment analysis identified genes differentially regulated
297 in both species as enriched in transcription factor-related GO terms, such as GO:0006355
298 "regulation of transcription, DNA-templated". This enrichment was further confirmed in a
299 separate test for enrichment of genes annotated as transcription factors in the GRASSIUS
300 database (Yilmaz et al., 2009). No non-transcription factor-related GO term showed significant
301 enrichment when compared to the population of gene pairs that were syntenically conserved
302 between both species. Comparison to the total population of annotated genes in maize or
303 sorghum showed many additional enrichments; however, this approach can produce misleading
304 results, as non-syntenic genes are enriched among genes without any functional annotation
305 (Schnable et al., 2012). We used MapMan (Usadel et al., 2009) to visualize the patterns of
306 expression within particular functional categories among DE2 genes as well as DE1 maize and
307 DE1 sorghum genes. As expected, genes related to cell wall growth, a marker for plant growth,
308 were downregulated in both species in the cold, including xyloglucosyl transferase
309 (Sobic.001g538000 and GRMZM2G388684) and leucine-rich repeat family protein

(Sobic.003g205600 and GRMZM2G333811) genes (Cui et al., 2005; Pearce, 2001; Tenhaken, 2014). Genes involved in lipid metabolism were upregulated in both species, including glycerol-3-phosphate acyltransferase 8 (Sobic.009g162000 and GRMZM2G166176), diacylglycerol kinase (Sobic.006g230400 and GRMZM2G106578), choline-phosphate cytidylyltransferase (Sobic.001g282900 and GRMZM2G132898), MGDG synthase (Sobic.004g334000 and GRMZM2G178892, Sobic.007g211900 and GRMZM2G141320), glycerophosphodiester phosphodiesterase (Sobic.007g190700 and GRMZM2G064962, Sobic.004g157300 and GRMZM2G018820) and FA elongation acyl-CoA ligase (Sobic.004g015400 and GRMZM2G120539) genes. This observation is consistent with the reported role of changes in membrane composition to avoid stiffening in the cold as an adaptive response to cold (Quinn, 1988; Singer and Nicolson, 1972). No consistent expression patterns of genes in particular metabolic processes (up- or downregulated) were observed among the DE1 maize or DE1 sorghum gene pairs.

The previously defined binding site for DREB/CBF transcription factors, which are induced in response to drought and cold stress (Muiño et al., 2016), showed significant enrichment in the proximal promoters of gene pairs in the DE2 category, as well as significant purification in the proximal promoters of gene pairs in the DE0 category (Supplemental Figure 5). As transcription factors are often associated with larger quantities of conserved noncoding sequences (CNS) (Freeling et al., 2007; Turco et al., 2013), we also investigated the number and quantity of conserved noncoding sequence associated with different classes of genes; however, no strong patterns were observed (Figure 6D). The use of conserved noncoding sequence data to identify regulatory sequence requires that the regulatory sequence be conserved between species. Given that many of the genes identified as responding to cold in either maize or sorghum appear to do so in a lineage-specific fashion, this requirement may not be satisfied in many cases. Various measurements of open chromatin have been shown to be good predictors of where regulatory sequences will be identified using CNS-based methods (Lai et al., 2017; Vera et al., 2014; Zhang et al., 2012), and unlike CNS-based methods, chromatin structure-based methods do not require that the same regulatory sequence be conserved across multiple species. We therefore examined the chromatin states in the promoters of genes with different patterns of cold-responsive regulation using a published dataset of MNase hypersensitive sites (HS) generated from maize

seedlings grown under non-stressed conditions (Rodgers-Melnick et al., 2016). Comparisons were made for maize DE0, Maize DE1, Sorghum DE1, DE2 and nonsyntenic genes at each of the six cold stress time points. Many nonsyntenic genes responded to cold; however, nonsyntenic genes as a whole showed little or no open chromatin (as defined by MNase HS) associated with their TSS (transcriptional start sites) or proximal promoters. Previous studies of other epigenetic marks have also concluded that the chromatin signatures of nonsyntenic genes in maize are more similar to those of intergenic sequences versus syntenic genes (Eichten et al., 2011). All categories of syntenic genes tended to have a peak of MNase sensitivity associated with their TSS and more open chromatin in their proximal promoters than nonsyntenic genes. Genes with conserved cold-responsive regulation (DE2) appear to have the greatest amount of open chromatin in their proximal promoters (Figure 7). Intriguingly, the maize copies of maize DE1 gene pairs exhibited stronger open chromatin signals that the maize copies of sorghum DE1 gene pairs, even though data on MNase hypersensitive sites came from seedlings grown under control conditions. The patterns reported above remained apparent when genes were divided into nine categories based on their relative expression level and Ka/Ks ratio, although statistical significance was reduced substantially as a result of the smaller number of genes included in each analysis (Supplemental Figure 6).

DISCUSSION

The above results indicate that there are roughly equivalent numbers of genes differentially expressed in response to cold compared to those reported from separate studies in each species (Chopra et al., 2015; Makarevitch et al., 2015). However, cross-species comparisons of the transcriptional regulation of the same genes in these two different species reveals that many cold-responsive patterns of regulation are not conserved between the two species. Correcting for the expected overlap across conserved genes based solely on the absolute genes number exhibiting cold-responsive transcriptional changes in each species further reduced the expected number of gene pairs where shared regulation resulted from the conservation of an ancestral pattern of cold-responsive transcriptional regulation. These data imply that gains or losses of cold-responsive regulation are relatively frequent in the grass tribe Andropogoneae. Genes that respond to cold in only a single lineage experience lower levels of purifying selection and are less likely to be annotated as transcription factor genes than genes that are cold-responsive in both lineages. It

13

should be noted that these results are based on data from a single accession of maize (B73) and a single accession of sorghum (BTx623). Evidence suggests that lower, but still significant, levels of divergence in transcriptional regulation in response to cold are present in different accessions of a single species (Makarevitch et al., 2015; Waters et al., 2017).

It appears that a relatively small core set of genes exhibit conserved responses to cold across the two species in this initial analysis, and functional analysis suggests that these genes are more likely to be present in pathways with logical links to cold stress (decreases in growth and cell wall biosynthesis, increases in lipid metabolism). Thus, we propose a model where a small core set of genes involved in the mechanisms by which panicoid grasses perceive and respond to cold stress are under functionally constrained cold-responsive transcriptional regulation, while a much larger set of genes can gain or lose cold-responsive transcriptional regulation in a neutral fashion or potentially as a result stabilizing selection, potentially through transposon-mediated mechanisms (Makarevitch et al., 2015; Naito et al., 2009). Consistent with this model, the genes with conserved cold-responsive gene regulation exhibited lower ratios of nonsynonymous to synonymous coding sequence substitutions than the other genes, which would imply their coding sequence is also subject to greater functional constraint. This model would also be consistent with the relatively high proportion of maize cold-responsive genes that exhibit variation in cold-responsive regulation across alleles (Waters et al., 2017).

We evaluated two different models for predicting conserved regulation across different expression levels and found that the multiplicative model was more effective at predicting orthologous gene pair expression than the additive model (Figure 5C, Supplemental Data Set 4). However, while this different was statistically significant, the additive model remained the better predictor for many gene pairs. While no obvious markers that distinguish genes where one model is the better predictor than the other were identified in this study, further study may identify additional molecular traits measured from the genome that can forecast which model is more appropriate for testing the expression pattern of a given gene across multiple related species.

The challenge of linking genes to functions based on expression evidence

The model above would predict than the observation of stress-responsive changes in transcript

14

abundance in a single species is not strong evidence that the associated gene plays a role in the response to that particular stress. While sequencing genomes and identifying genes are becoming more straightforward tasks, confidently assigning functional roles to newly identified genes remains challenging. Many genes in maize (35.1%) and sorghum (16.2%) are not associated with any GO annotations in the current release of Phytozome (v12). Many genes that do possess GO annotations are associated with only extremely broad annotation categories, such as protein binding or catalytic activity. "Guilt by association" studies using co-expression analysis are an intriguing method for assigning putative functional roles to some orphan or poorly annotated genes (Li et al., 2016; Schaefer et al., 2014). However, the use of these methods in a single species may also produce false positive annotations in the case of selectively neutral or stabilizing changes in gene regulation. It may prove to be the case that functionally constrained transcriptional responses are an effective method for identifying these links. Collecting parallel expression datasets in multiple species can be time consuming and costly. We therefore tested a number of alternative approaches to identifying functionally constrained cold-responsive transcriptional regulation. Early transcriptional responses to cold (30 minutes-3 hours) appeared to show greater conservation across species than later transcriptional responses. Regions of open chromatin detected through MNase HS (Rodgers-Melnick et al., 2016; Vera et al., 2014) were preferentially associated with genes that responded transcriptionally to cold stress in maize; however, this association was observed for genes with either conserved or lineage-specific patterns of cold-responsive regulation.

Importance of developing methods for cross-species comparisons of transcriptional regulation

Both modeling (Orr, 1998, 1999) and empirical studies (Chan et al., 2010; Studer et al., 2011) have found that genetic variants responsible for large, sudden changes in natural or artificial selection tend to have large, pleiotropic effects. In maize, distinct genetic architectures underlie traits that have been subjected to selection during domestication - one large-effect quantitative trait locus and many small modifiers - and traits that were not selected on during domestication - many small-effect quantitative trait loci (Wallace et al., 2014). This model was supported by recent work with an inter-subspecies cross of maize and its wild progenitor teosinte (*Zea mays* ssp. parviglumis). Looking at tassel morphology, distinctly genetic architectures were reported

15

for traits believed to have been under selection during domestication compared to those traits that were not (Xu et al., 2017). Developing effective approaches for comparing transcriptional regulation of conserved syntenic genes across related grass species has the potential to identify large-effect polymorphisms responsible for interspecies phenotypic variation in traits such as abiotic stress tolerance where substantial phenotypic variation exists between species (Figure 2D).

Here we have shown that by using synteny to identify pairs of conserved orthologs across related species, it is possible to identify species by treatment interactions, which signify changes in gene regulation across species (DROs), using a multiplicative model of gene regulation. The use of a multiplicative model was in turn supported by analysis of the regulation of duplicated maize genes within the same sample. By increasing the number of species sampled, it may soon be possible to define a consistent core set of genes subjected to functionally constrained regulation in response to cold across the grasses. Changes in the regulation of these core genes in specific lineages with different cold stress-response phenotypes would be useful candidates for the type of large-effect changes predicted to produce between-species phenotypic variation. However, the interpretation of such data must take into account that, unlike within-species studies of allelic variation in cold-responsive regulation, between-species analysis cannot distinguish cis-regulatory from trans-regulatory sources of variation in transcriptional responses.

METHODS

Plant growth and cold treatment

For maize and sorghum, the reference genotypes used for genome sequencing and assembly were B73 and BTx623, respectively. SNP calling using RNA-Seq data from B73 was used to verify that the plants used in this study came from the USA South clade of B73 accessions, i.e., those closest to the original reference genome (Liang and Schnable, 2016). Under the growing conditions employed, maize developed more quickly than sorghum, and sorghum seedlings twelve days after planting (DAP) were selected as being roughly developmentally equivalent to 10 DAP maize seedlings based on leaf number and morphology (Figure 2A). Planting dates were staggered so that all species reached this developmental time point simultaneously. For the original RNA-Seq presented in Figure 2A, seeds were planted in MetroMix 200 and grown in

16

greenhouse conditions under 13 hour day length in greenhouses at University of Nebraska-Lincoln's Beadle Center, with target conditions of 320 mol m^{-2} s^{-1}, high pressure sodium bulb, 13 hours/11 hours 29 ºC /23 ºC day/night and 60% relative humidity. Control plants were harvested directly from the greenhouse three hours before lights on. Plants subjected to cold stress treatment were moved to a cold treatment growth chamber, with 33 mol m^{-2} s^{-1}, metal halide grow bulb, 12 hours/12 hours 6ºC /6ºC day/night. Cold stressed plants were harvested three hours before lights on. Each sample consisted of pooled aboveground tissue from at least three seedlings. Each biological replicate was harvested from plants that were planted, grown, and harvested at a distinct and separate time from each other biological replicate. A total of three independent biological replicates where generated for this experiment. For the time course RNA-Seq data presented in Figure 4 and onward in the study, maize and sorghum were planted as above and grown in a Percival growth chamber (Percival model E-41L2) with target conditions of 111 mol m^{-2} s^{-1} light levels, 60% relative humidity, a 12 hour/12 hour day night cycle with a target temperature of 29 ºC during the day and 23 ºC at night. The onset of cold stress treatment was immediately before the end of daylight illumination, at which point half of the plants were moved to a second growth chamber with equivalent settings with the exception of a target temperature of 6 ºC both during the day and at night. Each sample represents a pool of all aboveground tissue from at least three seedlings. Samples were harvested from both the paired control and cold stress treatments at 0.5 hours, 1 hour, 3 hours, 6 hours, 16 hours, and 24 hours after the onset of cold stress. Biological replicates included both maize and sorghum plants that were offset in planting but stressed and harvested at the same time in the same growth chambers. A total of 3 independent biological replicates were generated for this experiment.

Definition of samples and biological replicates employed in this paper. **Sample**: each sample consists of RNA extracted from the pooled tissue of no less than 3 and no more than 5 separate plants planted and harvested on the same date and grown in the same growth chamber. All aboveground tissue was harvested from each plant included in a pool. **All aboveground tissue**: at the stage plants were harvested, all aboveground tissue included leaf blades, ligules, and leaf sheaths, but not apical meristems, stems, or roots. **Biological replicate**: each biological replicate consists of RNA extracted from pooled tissue harvested from plants of the same genotype planted and harvested on separate dates from any other biological replicate. **Paired replicate**:

biological replicates were paired across species, with tissue harvested on the same day from plants of each species growth in the same growth chamber.

CO_2 assimilation rate measurements

Plants were grown and cold treated as above, with the modification that in the case of sorghum, small plastic caps were placed over the seedlings to prevent the plants from becoming too tall to fit into the LiCor measurement chamber (approximately two inches). After 0, 1, or 3 days of cold treatment, the plants were allowed to recover in the greenhouse overnight. The following morning, CO_2 assimilation rates were measured using the Li-6400 portable photosystem unit under the following conditions: PAR 200 mol mol^{-1}, CO_2 at 400 mol mol-1 with flow at 400 mol mol^{-1} and humidity at greenhouse conditions. Whole plant readings were measured for sorghum, paspalum, Japanese millet (*Echincloa esculenta*), proso millet, and urochloa (*Urochloa fusca*) after covering their pots with clay and using the LiCor Arabidopsis chamber. Maize was measured using the leaf clamp attachment, which was consistently placed on the second leaf at a position 3 cm above the ligule. Leaf area was measured using the Li-3100c Area meter (Li-Cor). The accessions used for each species presented in Figure 1D included the following: paspalum: USDA PI 509022, Japanese millet: USDA PI 647850, proso millet: earlybird USDA PI 578073, urochloa: LBJWC-52, sorghum: BTx623, and maize: B73.

Identifying syntenic orthologs

Coding sequence data for primary transcripts of each annotated gene in the genome assemblies of 8 grass species, including maize and sorghum used in the analysis, were obtained from Phytozome 10.2. Similar sequences were identified using LASTZ (Harris, 2007), requiring an alignment spanning at least 50% of total sequence length and 70% sequence identity. In addition, the arguments -ambiguous=iupac, -notransition, and -seed=match12 were all set in each run. LASTZ output was converted to QuotaAlign's "RAW" format using a version of the blast to raw.py script that had been modified to take into account differences in output format between BLAST and LASTZ. The additional parameters -tandem Nmax=10 and -cscore=0.5 were specified when running this script.

RAW formatted data were processed using the core QuotaAlign algorithm with the parameters

18

-merge, and -Dm=20. -quota was set to 1:2 in comparisons to maize and 1:1 in all other comparisons. Pure QuotaAlign pan-grass syntenic gene sets were constructed using this dataset directly. Polished QuotaAlign pan-grass syntenic gene sets were constructed by first predicting the expected location for a given query gene in the target genome and then selecting the gene showing the greatest sequence similarity (as determined by lastz alignment score) within the window from 20 genes downstream of the predicted location to 20 genes upstream of the predicted location.

RNA-Seq data generation

RNA isolation and library construction followed the protocol described by Zhang et al. (Zhang et al., 2015). The number of reads generated per library is summarized in (Supplemental Data Set 1). Sequencing was conducted at Illumina Sequencing Genomics Resources Core Facility at Weill Cornell Medical College. Raw sequencing data are available through the NCBI (http://www.ncbi.nlm.nih.gov/bioproject) under accession number PRJNA343268 and PRJNA344653. Adapters were removed from raw sequence reads using cutadapt version 1.6 (Martin, 2011). RNA-Seq reads were mapped to genome assemblies downloaded from Phytozome: RefGen v3 (*Zea mays*), v3.1 (*Sorghum bicolor*). RNA-Seq reads from each species were aligned using GSNAP version 2014-12-29 (Wu and Nacu, 2010; Wu and Watanabe, 2005). Per-gene read counts were obtained using HTSeq v. 0.6.1 (Anders et al., 2014).

Identifying differentially expressed genes (DEGs)

Differentially expressed genes (DEGs) were identified using count data generated as described above and DESeq2 (version 1.14.0) (Love et al., 2014) based on a comparison of the treatment and control with adjP-value ≤ 0:05, meaning absolute $\log_2$ of fold change of between-treatment and control value ≥ 1. All expressed syntenic orthologous genes were classified into one of three categories. The three categories include genes that were classified as responding transcriptionally to cold in at least one species (DE1) (Figure 3A). The remaining category includes all expressed syntenic orthologous genes that were not classified as cold-responsive in either of the two species (DE0). The number of shared genes identified as differentially expressed in the two species (DE2) was tested relative to the expected overlap if there was no correlation in gene

19

regulation across species. For the time course RNA-Seq, analysis was conducted as above for all 36 possible pairwise comparisons of the six sorghum time points and six maize time points.

When estimating the true discovery proportion in analyses of DE2 genes (see Figure 3A, Figure 4B), it was necessary to calculate the number of DE2 genes expected under a null hypothesis of no conservation of gene regulation. This expected number of DE2 genes was calculated using the formula (percent of gene pairs DE in species 1)*(percent of gene pairs DE in species 2)*(total number of gene pairs analyzed was used). Total number of gene pairs was fixed at 15,232 syntenic orthologous gene pairs for maize1/sorghum comparisons and 9,554 for maize2/sorghum comparisons.

Estimating the power of DESeq2 in this dataset using simulated data

One-thousand genes were randomly sampled from the maize1/sorghum syntenic gene list in each repetition of the simulation. These selected genes included three replicates from both normal growth conditions (control) and one-day cold treatment (treatment). The geometric mean of each gene was calculated (adding 1 to the data to avoid 0 readings). A random sample from the uniform distribution on (5, 50) was used as the estimate of the true dispersion parameter. The simulated data for the non-differentially expressed genes were generated from a negative binomial distribution with the calculated geometric mean from the actual data and the sampled dispersion parameter. To generate the list of differentially expressed genes, the first 100 genes out of the 1000 sampled genes were selected with a treatment mean value equal to the geometric mean from the original data, whereas the mean value of the control was a multiple of the geometric mean (multiples of 2, 2.5, and 3 are reported). The calculated false discovery rate (FDR, ratio of number of false positives over total number of discoveries) and the power (ratio of true positives over the true number of differentially expressed genes) of the DESeq2 procedure are reported in Supplemental Data Set 2.

Evaluating the additive and multiplicative models of gene regulation

From the 5,257 duplicate genes retained from the maize WGD (Schnable et al., 2011) in each of the six time points in maize, gene pairs where both copies were classified as differentially expressed in response to cold were used to test both models. The expression pattern of the

maize1 gene under control and cold stress conditions plus the expression of the maize2 gene under control conditions was used to predict the expression of the maize2 gene under cold stress using both the additive and multiplicative models defined in Figure 5B. The distance between the prediction from the additive model and the observed value was defined as "a", the distance between the prediction from the multiplicative model and the observed value was defined as "b", and the predictions between the two models were defined as "c". In the relaxed case, gene pairs where the two models produced predictions that were closer to each other than either was to the observed expression value of the maize2 gene under cold stress were excluded. That is, if c<a and c<b, the multiplicative model works better than the additive model, while if b<a and b<c, the additive model works better than the other model. In the most stringent case, gene pairs where the two models produced predictions that were less than twice as large as the difference between the better model and the observed value were excluded (Supplemental Data Set 4). In other words, if b>2a and b>c, the multiplicative model was considered to be the better model; if c>2a and c>b, the additive model was considered to be the better model. Analyses were also conducted reciprocally using data from control and cold stress conditions in maize2 plus data from maize1 under control conditions to predict the expression of the maize1 gene under cold stress conditions.

Identifying differentially regulated orthologs (DRO)

Differentially regulated orthologs were identified using count data generated as described above and an interaction term for species (maize or sorghum) and treatment (cold or control) in DESeq2 (Love et al., 2014). Species (maize and sorghum) and condition (cold and control) were considered to be two factors for design in this analysis. Simulated data for comparably regulated orthologs (CROs) generated using additive and multiplicative models were used to confirm that this approach did not classify simulated CROs based on the multiplicative model as having significant species-by-treatment interactions. The formula used was: design _ condition + genotype + condition: genotype. Maize sorghum gene pairs with an interaction adjP-value ≤ 0:001 were classified as DROs, those with interaction adjP-value ≥ 0:05 were classified as CROs, and those with intermediate p-values were disregarded (Yoav and Yosef, 1995). The decision was made to retain an ambiguous case of gene pairs with interaction p-values too high to be classified as DROs but too significant to be classified as CROs rather than increase the number

of classification errors by forcing all gene pairs to be assigned to one category or the other.

Calculating Ka/Ks values

"Primary Transcript only" coding sequences for maize (v6a), sorghum (v3.1), and setaria (v2.2) were retrieved from Phytozome version 12.0. The gene model annotations v6a for maize were annotated onto the B73 RefGen v3 pseudomolecules. Coding sequences were translated to protein sequences and aligned using Kalign version 2.04 (Lassmann and Sonnhammer, 2005). The protein alignment was used as a guide to create a codon level alignment of coding sequences. The codon alignment was supplied to PAML (version 4.09) (Yang, 2007). Synonymous and nonsynonymous substitution rates were calculated independently for each branch of the tree. When both a maize1 and maize2 gene copies were present for the same syntenic gene group, alignment and substitution rate calculations were conducted separately for the maize1 gene and its syntenic orthologs in sorghum and setaria and for the maize2 and the same syntenic orthologous genes. To eliminate genes with extreme Ka/Ks ratios resulting from very low numbers of synonymous substitutions, only Ka/Ks ratios from genes with an estimated synonymous substitution rate greater than or equal to 0.05 (approximately 1/2 the median Ks ratio observed between maize and the most common recent ancestor of maize and sorghum) were considered.

MNase hypersensitive site analysis

Intervals defined as MNase hypersenstive sites (MNase HS) were taken from (Rodgers-Melnick et al., 2016). The same TSS was used for MNase and RNA-Seq analysis. Average coverage of MNase HS was calculated on a per-base basis from 1 kb upstream of the annotated TSS to 1 kb downstream of the TSS. When multiple transcripts with different TSS were present, the transcript with the earliest TSS was selected for analysis.

Identifying Conserved Noncoding Sequences (CNS)

CNSs were identified using the CNS Discovery Pipeline 3.0 (CDP) (Turco et al., 2013) with some modifications. Specifically, the built-in syntenic gene identification pipeline from the CDP was replaced with the previously defined syntenic gene list described above. Functions for finding local duplicates and comparing CNSs to Arabidopsis proteins and RNA were omitted.

CNSs were identified between the region 12 kb upstream and 12 kb downstream using a word size of 15 bp. CNSs with bit scores for each gene pair < 29.5 were removed following the same scoring parameter settings outlined in the original software pipeline.

Transcription Factor Enrichment Calculation

Transcription Factor (TF) enrichment was calculated using the maize transcription factor list from GRASSIUS (Yilmaz et al., 2009).

GO enrichment analysis

Gene ontology (GO) analysis was performed using GOATOOLS (Haibao et al., 2015) and functional additions associated with the sorghum v3.1 sorghum gene model and maize RefGen-v3 maize gene model annotations.

Pathway analysis

Pathway analysis was conducted using the MapMan software package (http://mapman.gabipd.org/web/guest) (Usadel et al., 2009).

Accession numbers

GeneIDs for all syntenic gene sets and the final syntenic gene list used in this study are posted at figShare (http://dx.doi.org/10.6084/m9.figshare.3113488.v1).

Adapter sequences used for library construction and for adapter trimming are those provided in Illumina TruSeq Library Prep Pooling Guide, with sequences reported on page 5 of the user manual.

Supplemental Data

Supplemental Figure 1. Coding sequence similarity among syntenic genes in sorghum, maize1, and maize2.

Supplemental Figure 2. Representative sample of cold stressed seedling phenotypes.

Supplemental Figure 3. Individual examples of genes in each of six possible DRO/DEG classification categories.

23

682 **Supplemental Figure 4.** Comparison of Ka/Ks ratio and expression level for genes
683 grouped based on expression classification model.
684 **Supplemental Figure 5.** Frequency of known CBF binding motifs within the 1 kb
685 proximal promoters of maize and sorghum.
686 **Supplemental Figure 6.** Relationship between gene pair expression pattern in maize and
687 sorghum after subdividing genes based on Ka/Ks ratio and expression tertile.
688
689 **Supplemental Data Set 1.** Number of sequenced and aligned reads per library.
690
691 **Supplemental Data Set 2.** Estimates of power and FDR for DESeq2.
692
693 **Supplemental Data Set 3.** ANOVA and DESeq2 tests for DROs using simulated data.
694
695 **Supplemental Data Set 4.** Accuracy of additive and multiplicative expression models across
696 maize duplicate gene pairs.
697
698 **ACKNOWLEDGMENTS**
699 This material is based upon work supported by the National Institute of Food and Agriculture,
700 U.S. Department of Agriculture under award number 016-67013-24613 to RLR and JCS. This
701 material is based upon work supported by the National Science Foundation under Grant No.
702 OIA-1557417. In addition, this work was supported by startup funding from the University of
703 Nebraska Lincoln to RLR, YQ, and JCS.
704
705 **AUTHOR CONTRIBUTIONS**
706
707 JCS and RLR conceived the project and designed the studies; YZ, DWN, DC, and ZL performed
708 the research; YZ and YQ analyzed the data; YZ, JCS, and RLR wrote the paper. All authors
709 reviewed the manuscript.
710
711 **REFERENCES**
712 **Anders, S., Pyl, P.T., and Huber, W.** (2014). HTSeq-a python framework to work with high-
713 throughput sequencing data. Bioinformatics. **31(2)**: 166-169.

Yoav, B., Y., Yosef, H., (1995)**.** Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Ser B Methodol. **57:** 289-300.

Bennetzen, J. L., and Freeling, M. (1993). Grasses as a single genetic system: genome composition, collinearity and compatibility. Trends Genet. **9(8):** 259-261.

Burkart-Waco, D., Ngo, K., Lieberman, M., and Comai, L. (2015). Perturbation of parentally biased gene expression during interspecific hybridization. PLoS One. **10(2):** e0117293.

Chan, Y.F., Marks, M.E., Jones, F.C., Villarreal, G., Shapiro, M.D., Brady, S.D., Southwick, A.M., Absher, D.M., Grimwood, J., Schmutz, J., **Myers, R.M., Petrov D., Jónsson B., Schluter, D., Bell M.A., Kingsley, D.M.** (2010). Adaptive evolution of pelvic reduction in sticklebacks by recurrent deletion of a pitx1 enhancer. Science. **327(5963):** 302-305.

Chinnusamy, V., Zhu, J., and Zhu, J.K. (2007). Cold stress regulation of gene expression in plants. Trend Plant Sci. **12(10):** 444-451.

Chopra, R., Burow, G., Hayes, C., Emendack,Y., Xin, Z., and Burke, J. (2015). Transcriptome profiling and validation of gene based single nucleotide polymorphisms (SNPs) in sorghum genotypes with contrasting responses to cold stress. BMC genomics. **16:** 1040.

Chopra, R., Burow, G., Burke, J. J., Gladman, N., and Xin, Z. (2017). Genome-wide association analysis of seedling traits in diverse sorghum germplasm under thermal stress. BMC Plant Biol. **17:** 12.

Christin, P.A., Salamin, N., Savolainen, V., Duvall, M.R., and Besnard, G. (2007). C_4 photosynthesis evolved in grasses via parallel adaptive genetic changes. Current Biology, **17(14):** 1241-1247.

Cooper, G.M., Stone, E.A., Asimenos, G., Green, E.D., Batzoglou, S., and Sidow, A. (2005). Distribution and intensity of constraint in mammalian genomic sequence. Genome Res. **15(7):** 901-913.

Cui, S., Huang, F., Wang, J., Ma, X., Cheng, Y., and Liu, J. (2005). A proteomic analysis of cold stress responses in rice seedlings. Proteomics. **5(12):** 3162-3172.

Davidson, R. M., Gowda, M., Moghe, G., Lin, H., Vaillancourt, B., Shiu S.-H, Jiang N., and Buell, C.R. (2012). Comparative transcriptomics of three poaceae species reveals patterns of gene expression evolution. Plant J. **71(3):** 492-502.

De Wet J. (1978). Systematics and evolution of sorghum sect. sorghum (gramineae). Am J Bot. **65(4):** 477-484.

Eichten, S.R., Swanson-Wagner, R.A., Schnable, J.C, Waters, A.J., Hermanson, P.J., Liu, S., Yeh, CT., Jia, Y., Gendler, K., Freeling, M., Schnable, P.S., Vaughn, M.W., Springer, N.M. (2011). Heritable epigenetic variation among maize inbreds. PLoS Genet. **7(11):** e1002372.

Ercoli, L., Mariotti, M., Masoni, A., and Arduini, I. (2004). Growth responses of sorghum plants to chilling temperature and duration of exposure. Eur J Agron. **21(1):** 93-103.

Fiedler, K., Bekele, W. A., Matschegewski, C., Snowdon, R., Wieckhorst, S., Zacharias, A., and Uptmoor, R. (2016). Cold tolerance during juvenile development in sorghum: a comparative analysis by genomewide association and linkage mapping. Plant Breeding., **135(5):** 598-606.

Freeling, M., Rapaka, L., Lyons, E., Pedersen, B., and Thomas, B.C. (2007). G-boxes, bigfoot genes, and environmental response: characterization of intragenomic conserved noncoding sequences in arabidopsis. Plant Cell. **19(5):** 1441-1457.

Freeling, M., Woodhouse, M.R., Subramaniam, S., Turco, G., Lisch, D., and Schnable, J.C. (2012). Fractionation mutagenesis and similar consequences of mechanisms removing dispensable or less-expressed dna in plants. Curr Opin Plant Biol. **15(2):** 131-139.

Haibao, T., Klopfenstein, D., Pedersen, B., Flick, P., Sato, K., Ramirez, F., Yunes, J., and Mungall, C. (2015). Goatools: Tools for gene ontology. Zenodo.

Harris R.S. (2007). Improved pairwise alignment of genomic DNA. Ph.D. thesis Pennsylvania State University.

Hetherington, S. E., He, J., and Smillie, R.M. (1989). Photoinhibition at low temperature in chilling-sensitive and-resistant plants. Plant Physiol. **90(4):** 1609-1615.

26

Hollister, J.D. and Gaut, B.S. (2009). Epigenetic silencing of transposable elements: a trade-off between reduced transposition and deleterious effects on neighboring gene expression. Genome Res. **19(8):** 1419-1428.

Hollister J.D., Smith L.M., Guo, Y.-L., Ott, F., Weigel, D., and Gaut, B.S. (2011). Transposable elements and small RNAs contribute to gene expression divergence between *Arabidopsis thalian*a and *Arabidopsis lyrata.* Proc. Natl. Acad. Sci. U.S.A. **108(6):** 2322-2327.

Lai, X., Behera, S., Liang, Z., Lu, Y., Deogun, J. S., and Schnable, J. C. (2017). Stag-CNS: An order-aware conserved non-coding sequences discovery tool for arbitrary numbers of species. Mol Plant.

Lassmann, T., and Sonnhammer, E.L. (2005). Kalign-an accurate and fast multiple sequence alignment algorithm. BMC bioinformatics. **6(1):** 298.

Li, L., Briskine, R., Schaefer, R., Schnable, P.S., Myers, C.L., Flagel, L.E., Springer, N.M., and Muehlbauer, G.J. (2016). Co- expression network analysis of duplicate genes in maize (*Zea mays L.*) reveals no subgenome bias. BMC genomics, **17(1):** 875.

Liang, Z. and Schnable, J.C. (2016). RNA-Seq based analysis of population structure within the maize inbred B73. PloS One. **11(6):** e0157942.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-Seq data with DESeq2. Genome Biol. **15(12):** 550.

Lovell, J. T., Schwartz, S., Lowry, D. B., Shakirov, E. V., Bonnette, J. E., Weng, X., Wang, M., Johnson, J., Sreedasyam, A., Plott, C., et al. (2016). Drought responsive gene expression regulatory divergence between upland and lowland ecotypes of a perennial c4 grass. Genome Res. **26(4):** 510-518.

Makarevitch, I., Waters, A.J., West, P.T., Stitzer, M., Hirsch, C.N., Ross-Ibarra, J., and Springer, N.M (2015). Transposable elements contribute to activation of maize genes in response to abiotic stress. PLoS Genet. **11(1)**: e1004915.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal. **17(1):**10-12.

27

Muiño, J.M., de Bruijn, S., Pajoro, A., Geuten, K., Vingron, M., Angenent, G.C., and Kaufmann, K. (2016). Evolution of dna-binding sites of a floral master regulatory transcription factor. Mol Biol Evol. **33(1):** 185-200.

Naito, K., Zhang, F., Tsukiyama, T., Saito, H., Hancock, C.N., Richardson, A.O., Okumoto, Y., Tanisaka, T., and Wessler, S.R. (2009). Unexpected consequences of a sudden and massive transposon amplification on rice gene expression. Nature. **461(7267):** 1130-1134.

Ng, P.C and Henikoff, S. (2001). Predicting deleterious amino acid substitutions. Genome Res. **11(5):** 863-874.

Olsen, J., McMahon, C., and Hammer, G. (1993). Prediction of sweet corn phenology in subtropical environments. Agron J. **85(2):** 410-415.

Orr, H.A. (1998). The population genetics of adaptation: the distribution of factors fixed during adaptive evolution. Evolution. **52(4):** 935-949.

Orr, H.A., (1999). The evolutionary genetics of adaptation: a simulation study. Genet Res. **74(03):** 207-214.

Paschold, A., Larson, N.B., Marcon, C., Schnable, J.C., Yeh C.-T., Lanz, C., Nettleton, D., Piepho, H.-P., Schnable, P.S., and Hochholdinger, F. (2014). Nonsyntenic genes drive highly dynamic complementation of gene expression in maize hybrids. Plant Cell. **26(10):** 3939-3948.

Paterson, A.H., Bowers J.E., Bruggmann, R., Dubchak, I., Grimwood, J., Gundlach, H., Haberer, G., Hellsten, U., Mitros, T., Poliakov, A., et al. (2009). The Sorghum bicolor genome and the diversification of grasses. Nature. **457(7229):** 551-556.

Pearce, R. S., (2001). Plant freezing and damage. Annals of Botany, **87(4):** 417-424.

Priest, H.D., Fox, S.E., Rowley, E.R., Murray, J.R., Michael, T.P., and Mockler, T.C. (2014). Analysis of global gene expression in *brachypodium distachyon* reveals extensive network plasticity in response to abiotic stress. PLoS One. **9(1):** e87499.

Quinn, P. J., (1988). Effects of temperature on cell membranes. Symp Soc Exp Biol. **42:** 237-258.

Reva, B., Antipin, Y., and Sander, C. (2011). Predicting the functional impact of protein mutations: application to cancer genomics. Nucleic Acids Res. **39(17):** e118.

Revilla, P., Rodríguez, V.M., Ordás, A., Rincent, R., Charcosset, A., Giauffret, C., Melchinger, A.E., Schön, C.-C., Bauer, E., Altmann, T., et al. (2016). Association mapping for cold tolerance in two large maize inbred panels. BMC Plant Biol. **16:** 127.

Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W., and Smyth G.K. (2015). Limma powers differential expression analyses for RNA-Sequencing and microarray studies. Nucleic Acids Res. **43 (7):** e47

Robinson M. D, McCarthy D. J, and Smyth G. K 2010. edger: a bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics, **26(1):** 139-140.

Rodgers-Melnick, E., Vera D.L., Bass, H.W., and Buckler, E.S. (2016). Open chromatin reveals the functional maize genome. Proc. Natl. Acad. Sci. U.S.A. **113(22):** E3177-E3184.

Sánchez, B., Rasmussen, A., and Porter, J.R. (2014). Temperatures and the growth and development of maize and rice: a review. Global Change Biol. **20(2):** 408-417.

Schaefer, R.J., Briskine, R., Springer, N.M., and Myers, C.L. (2014). Discovering functional modules across diverse maize transcriptomes using cob, the co-expression browser. PloS one, **9(6):** e99193.

Schnable, J. C., (2015). Genome evolution in maize: From genomes back to genes. Annu Rev Plant Biol. **66:** 329-343.

Schnable, J.C, Springer, N.M., and Freeling, M. (2011). Differentiation of the maize subgenomes by genome dominance and both ancient and ongoing gene loss. Proc Natl Acad Sci U.S.A. **108(10):** 4069-4074.

Schnable J.C., Freeling, M., and Lyons, E. (2012). Genome-wide analysis of syntenic gene deletion in the grasses. Genome Biol Evol. **4(3):** 265-277.

Schnable, P.S., Ware, D., Fulton, R.S., Stein, J.C., Wei, F., Pasternak, S., Liang, C., Zhang, J., Fulton, L., Graves, T.A., et al. 2009. The B73 maize genome: complexity, diversity, and dynamics. Science. **326(5956):** 1112-1115.

Schurch, N. J., Schofield, P., Gierlinski, M., Cole, C., Sherstnev, A., Singh, V., Wrobel, N., Gharbi, K., Simpson, G. G., Owen-Hughes, T., et al. (2016). How many biological replicates are needed in an RNA-Seq experiment and which differential expression tool

should you use? RNA. **22(6):** 839-851.

Shantz, H. (1954). The place of grasslands in the earth's cover. Ecology. **35(2):** 143-145.

Shaykewich, C. (1995). An appraisal of cereal crop phenology modelling. Can J Plant Sci. **75(2):** 329-341.

Studer, A., Zhao, Q., Ross-Ibarra, J., and Doebley, J. (2011). Identification of a functional transposon insertion in the maize domestication gene tb1. Nature Genet. **43(11):** 1160-1163.

Swigoňová, Z., Lai, J., Ma, J., Ramakrishna, W., Llaca, V., Bennetzen, J.L., and Messing, J. (2004). Close split of sorghum and maize genome progenitors. Genome Res. **14(10a):** 1916-1923.

Tenhaken, R., (2014). Cell wall remodeling under abiotic stress. Front. Plant Sci., **5**. doi: 10.3389/fpls.2014.00771.

Tiwari, S., Krishnamurthy, S., Kumar, V., Singh, B., Rao, A., SV, A.M., Rai,V., Singh, A.K., and Singh, N.K. (2016). Mapping qtls for salt tolerance in rice (**Oryza sativa L.**) by bulked segregant analysis of recombinant inbred lines using 50k snp chip. PloS One., **11(4):** e0153610.

Torres-Oliva, M., Almudi, I., McGregor, A.P., and Posnien, N. (2016). A robust (re-) annotation approach to generate unbiased mapping references for RNA-Seq-based analyses of differential expression across closely related species. BMC genomics, **17(1)**: 392.

Turco, G., Schnable, J.C., Pedersen, B., and Freeling, M. (2013). Automated conserved non-coding sequence (CNS) discovery reveals differences in gene content and promoter evolution among grasses. Front Plant Sci. **4**: 170.

Usadel, B., Poree, F., Nagel, A., Lohse, M., CZEDIK-EYSENBERG A., and Stitt, M. (2009). A guide to using mapman to visualize and compare omics data in plants: a case study in the crop species, maize. Plant Cell Environ. **32(9):** 1211-1229.

van Heerwaarden, J., Doebley, J., Briggs, W.H., Glaubitz, J.C., Goodman, M.M., Gonzalez, J.d.J.S., and Ross-Ibarra, J. (2011). Genetic signals of origin, spread, and introgression in a large sample of maize landraces. Proc Natl Acad Sci U.S.A. **108(3):** 1088-1092.

Vera, D.L., Madzima, T.F., Labonne, J.D., Alam M.P., Hoffman, G.G., Girimurugan, S., Zhang, J., McGinnis, K.M., Dennis, J.H., and Bass, H.W. (2014). Differential nuclease sensitivity profiling of chromatin reveals biochemical footprints coupled to gene expression and functional dna elements in maize. Plant Cell. **26(10):** 3883-3893.

Wallace, J., Larsson, S., and Buckler, E. (2014). Entering the second century of maize quantitative genetics. Heredity, **112(1):** 30-38.

Wang, L., Czedik-Eysenberg, A., Mertz, R. A., Si, Y., Tohge, T., Nunes-Nesi, A., Arrivault, S., Dedow, L. K., Bryant, D. W., Zhou, W., et al. (2014). Comparative analyses of C_4 and C_3 photosynthesis in developing leaves of maize and rice. Nature Biotechnol. **32(11):** 1158-1165.

Waters, A.J, Makarevitch, I., Noshay, J., Burghardt, L.T, Hirsch, C.N, Hirsch, C.D., and Springer, N.M. (2017). Natural variation for gene expression responses to abiotic stress in maize. Plant J. **89:** 706-717.

Wendorf, F., Close, A.E., Schild, R., Wasylikowa, K., Housley, R.A., Harlan, J.R., and Krolik, H. (1992). Saharan exploitation of plants 8,000 years bp. Nature, **359(6397):** 721-724.

Wu, T.D and Nacu, S. (2010). Fast and snp-tolerant detection of complex variants and splicing in short reads. Bioinformatics. **26(7):** 873-881.

Wu, T.D and Watanabe, C.K (2005). Gmap: a genomic mapping and alignment program for mRNA and EST sequences. Bioinformatics. **21(9):** 1859-1875.

Xu, G., Wang, X., Huang, C., Xu, D., Li, D., Tian, J., Chen, Q., Wang, C., Liang, Y., Wu, Y., et al. (2017). Complex genetic architecture underlies maize tassel domestication. New Phytol. **214:** 852-864.

Yang, Z. (2007). Paml 4: phylogenetic analysis by maximum likelihood. Mol. Biol. Evol. **24(8):** 1586-1591.

Yilmaz, A., Nishiyama, M.Y., Fuentes, B.G., Souza, G.M., Janies, D., Gray, J., and Grotewold, E. (2009). Grassius: a platform for comparative regulatory genomics across the grasses. Plant Physiol. **149(1):** 171-180.

Zhang, W., Wu, Y., Schnable, J. C., Zeng, Z., Freeling, M., Crawford, G. E., and Jiang, J. (2012). High-resolution mapping of open chromatin in the rice genome. Genome Res.

22(1): 151-162.

Zhang, Y., Ding, Z., Ma, F., Chauhan, R.D., Allen, D.K., Brutnell, T.P., Wang, W., Peng, M., and Li, P. (2015). Transcriptional response to petiole heat girdling in cassava. Sci Rep. **5**: 8414.

FIGURE LEGENDS

Figure 1. Gene level and expression level conservation between sorghum, maize1, and maize2. (A) The overlap between syntenic orthologous gene pairs conserved between maize1/sorghum and maize2/sorghum. **(B)** Comparison of average control condition expression levels (log_2 transformed FPKM) for either maize1/sorghum or maize2/sorghum gene pairs. (To improve readability, a random sample of 1/3 of all gene pairs is displayed for each category.)

Figure 2. Effects of cold stress on maize, sorghum, and related species. Representative seedling phenotypes for maize and sorghum. **(A)** Control conditions; **(B)** 24 hours of stress at 6°C; **(C)** 14 days at 6°C and two days recovery under greenhouse conditions. **(D)** Normalized relative CO_2 assimilation rates for six panicoid grass species with differing degrees of sensitivity or tolerance to cold stress. Individual data points were jittered (adding random noise to data in order to prevent over-plotting in statistical graphs) on the x-axis to avoid overlap and improve readability.

Figure 3. Combined DEG analysis of maize and sorghum. (A) An illustration of the DEG-based gene pair classification model and a comparison of expected and observed values for gene pairs classified as differentially expressed in response to cold in zero, one, or both species. Expected distributions were calculated based on a null hypothesis of no correlation in gene regulation between maize and sorghum (see Methods). DE0: gene pairs classified as differentially expressed in response to cold in neither species; DE1: gene pairs classified as differentially expressed in response to cold in one species but not the other; DE2: gene pairs classified as differentially expressed in response to cold in both species. (Observed number of gene pairs in maize1/sorghum: DE1 maize = 850, DE2 = 836, DE1 sorghum = 1,507, DE0 = 12,038; Observed number of gene pairs in maize2/sorghum: DE1 maize = 508, DE2 = 460, DE1 sorghum = 986, DE0 = 7,599; Expected number of gene pairs in maize1/sorghum: DE1 maize =

32

1,427, DE2 = 259, DE1 sorghum = 2,084, DE0 = 11,461; Expected number of gene pairs in maize2/sorghum: DE1 maize = 822, DE2 = 146, DE1 sorghum = 1,300, DE0 = 7,285). **(B)** Comparison of fold change in gene expression between the treatment and control groups for pairs of orthologous genes in maize and sorghum. Log_2 transformed treatment/control expression ratios are shown.

Figure 4. Patterns of gene expression across a cold-stress time series in maize and sorghum. **(A)** Changes in classification of individual gene pairs as DE0, DE1 maize, DE1 sorghum, and DE2 across adjacent time points. (**B)** The proportion of genes identified as differentially expressed in both species in excess of the number of gene pairs expected in this category in the absence of either conservation of gene regulation or parallel evolution of gene regulation. True discovery proportion is defined as (Observed Positives - Estimated False Positives)/Observed Positives. The expected number false positive DE2 gene pairs was calculated from the proportion of all genes classified as DEGs in maize and sorghum using the null model described in Figure 3A.

Figure 5. Conceptual illustration of the differentially regulated ortholog model. (A) Illustration of the different classification outcomes that can be produced for a given gene pair using both a DEG-based analysis (testing whether the expression pattern of each gene changes significantly between conditions) and a DRO-based analysis (testing whether the pattern across the two conditions is significantly different between copies of the same gene in both species). **(B)** Two models, additive and multiplicative, for predicting what a conserved pattern of gene regulation should look like when the underlying level of expression changes. **(C)** Relationship between prediction error (log_{10} transformed) for expression under cold stress using a multiplicative model to predict expression between maize1/maize2 gene pairs or an additive model to predict expression between maize1/maize2 gene pairs. Maize1: Predictions for the expression pattern of maize2 genes using data from their maize1 homeologs; Maize2: Predictions for the expression pattern of maize1 genes using data from their maize2 homeologs. Blue dots mark cases where the additive model was the better predictor; red dots mark cases where the multiplicative model was the better predictor.

33

Figure 6. Characteristics of genes in different DEG groups at different timepoints. (A) The proportion of gene pairs classified as DROs between maize and sorghum in different DEG groups at each of the six time points examined. **(B)-(C)** Median ratios of non-synonymous substitutions to synonymous substations in coding sequences for maize and sorghum for gene pairs classified as DE0, DE1, or DE2 at each of six time points. Time points where there is a statistically significant difference in Ka/Ks ratio between DE2 and any of the other three categories are marked with either + (if $p<0.05$) or ++ (if $P<0.01$). Color of the + indicates the category to which DE2 is being compared. Time points where there is a statistically significant difference in Ka/Ks ratio between DE0 and either DE1 maize or DE1 sorghum categories are marked with either * (if $p<0.05$) or ** (if $P<0.01$). Color of the * indicates the category to which DE0 is being compared. (B)-(C) Enrichment of genes annotated as transcription factor genes among DE2 gene pairs relative to all syntenic gene pairs indicated by the black line and the right hand axis. Double white triangles mark time points where the enrichment is statistically significant ($p < 0.01$). **(D)** Frequency of CNS within the promoters of genes classified as DE0, DE1 maize, DE1 sorghum, DE2, DRO, or CRO at each of the six time points. Black lines within the box plot mark the average number of CNS per gene for each category.

Figure 7. **Chromatin patterns associated with different groups of genes in maize and sorghum.** Patterns of MNase hypersensitive regions around the transcriptional start sites of genes classified based on their pattern of gene regulation in the 24 hour stress time point. Maize1 sorghum gene pairs and maize2 sorghum gene pairs were aggregated to increase statistical power. The lighter band around the DE2 line indicates a two standard deviation confidence interval. Black bars at the bottom of the graph indicate individual base pair positions where the amount of open chromatin associated with DE2 genes is significantly different from that of each of the other four categories displayed with a p-value < 0.01 for each comparison. Pairwise comparisons were performed using a Fisher Exact Test.

34

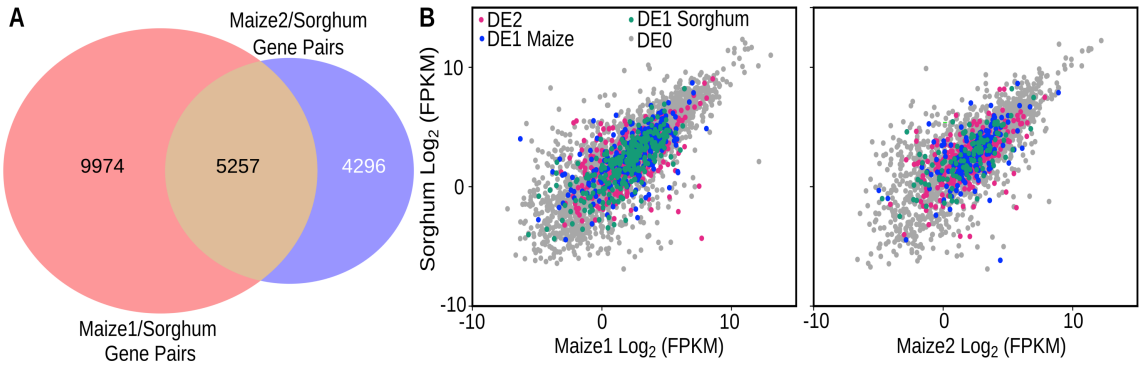

Figure 1. Gene level and expression level conservation between sorghum, maize1, and maize2. (A) The overlap between syntenic orthologous gene pairs conserved between maize1/sorghum and maize2/sorghum. **(B)** Comparison of average control condition expression levels ($\log_2$ transformed FPKM) for either maize1/sorghum or maize2/sorghum gene pairs. (To improve readability, a random sample of 1/3 of all gene pairs is displayed for each category.) Please use subscript for "$\log_2$" in the figure.

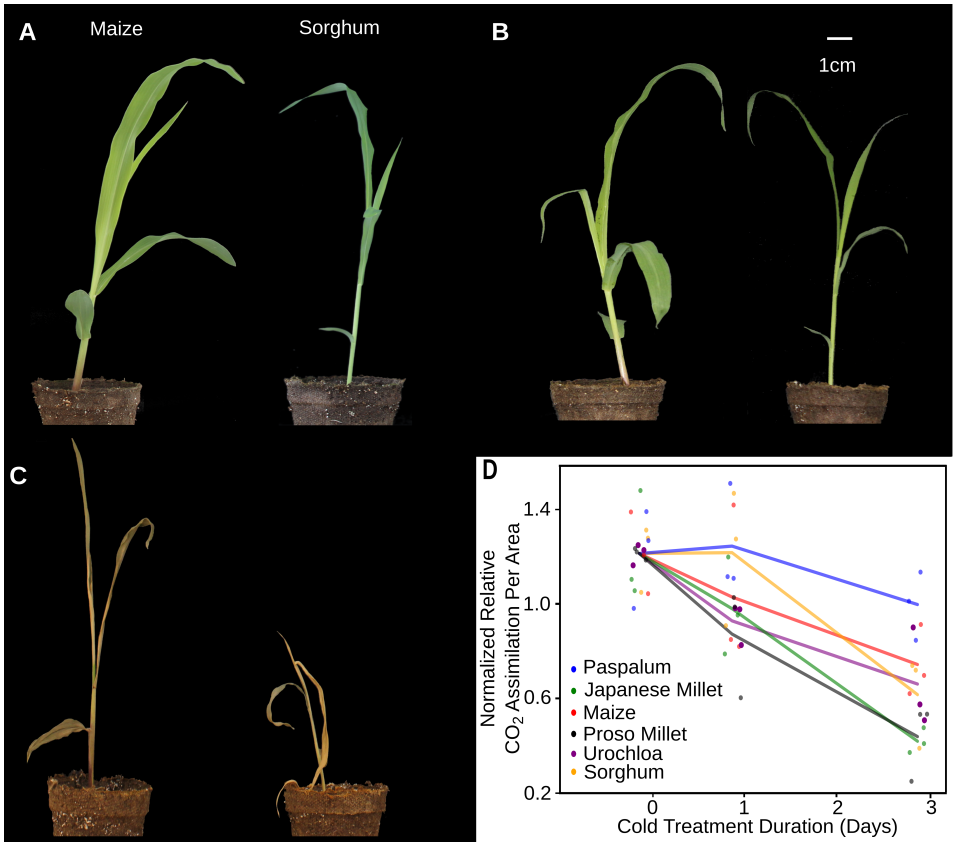

Figure 2. Effects of cold stress on maize, sorghum, and related species. Representative seedling phenotypes for maize and sorghum. **(A)** Control conditions; **(B)** 24 hours of stress at 6°C; **(C)** 14 days at 6°C and two days recovery under greenhouse conditions. **(D)** Normalized relative CO_2 assimilation rates for six panicoid grass species with differing degrees of sensitivity or tolerance to cold stress. Individual data points were jittered (adding random noise to data in order to prevent over-plotting in statistical graphs) on the x-axis to avoid overlap and improve readability.

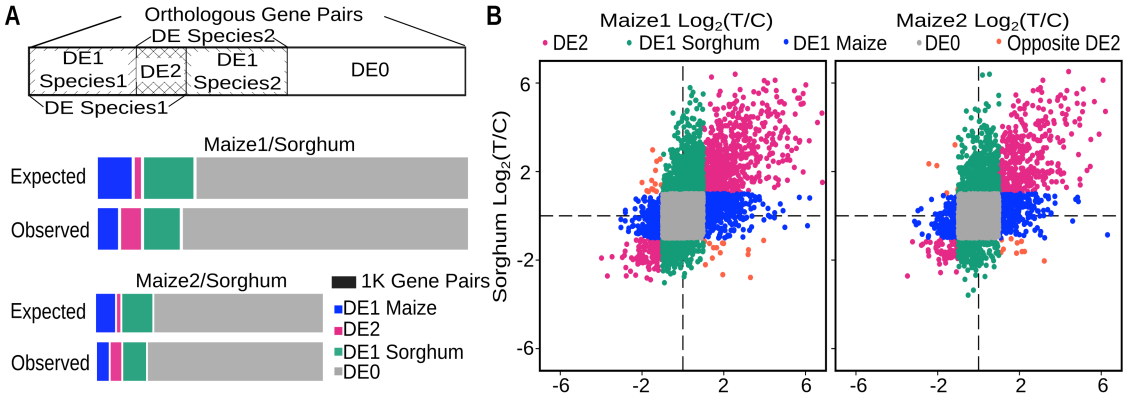

Figure 3. Combined DEG analysis of maize and sorghum. (A) An illustration of the DEG-based gene pair classification model and a comparison of expected and observed values for gene pairs classified as differentially expressed in response to cold in zero, one, or both species. Expected distributions were calculated based on a null hypothesis of no correlation in gene regulation between maize and sorghum (see Methods). DE0: gene pairs classified as differentially expressed in response to cold in neither species; DE1: gene pairs classified as differentially expressed in response to cold in one species but not the other; DE2: gene pairs classified as differentially expressed in response to cold in both species. (Observed number of gene pairs in maize1/sorghum: DE1 maize = 850, DE2 = 836, DE1 sorghum = 1,507, DE0 = 12,038; Observed number of gene pairs in maize2/sorghum: DE1 maize = 508, DE2 = 460, DE1 sorghum = 986, DE0 = 7,599; Expected number of gene pairs in maize1/sorghum: DE1 maize = 1,427, DE2 = 259, DE1 sorghum = 2,084, DE0 = 11,461; Expected number of gene pairs in maize2/sorghum: DE1 maize = 822, DE2 = 146, DE1 sorghum = 1,300, DE0 = 7,285). **(B)** Comparison of fold change in gene expression between the treatment and control groups for pairs of orthologous genes in maize and sorghum. Log_2 transformed treatment/control expression ratios are shown.

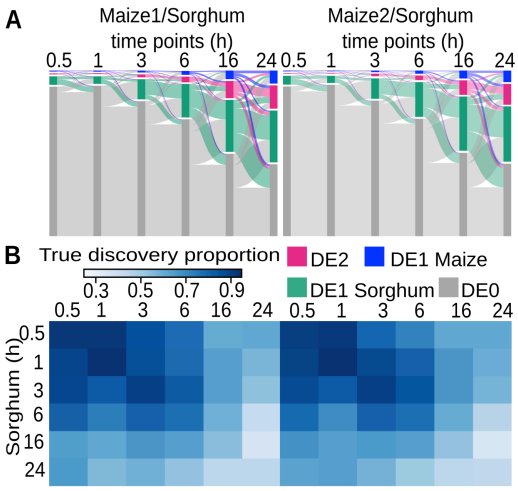

Figure 4. Patterns of gene expression across a cold stress time series in maize and sorghum. (A) Changes in classification of individual gene pairs as DE0, DE1 maize, DE1 sorghum, and DE2 across adjacent time points. **(B)** The proportion of genes identified as differentially expressed in both species in excess of the number of gene pairs expected in this category in the absence of either conservation of gene regulation or parallel evolution of gene regulation. True discovery proportion is defined as (Observed Positives - Estimated False Positives)/Observed Positives. The expected number false positive DE2 gene pairs was calculated from the proportion of all genes classified as DEGs in maize and sorghum using the null model described in Figure 3A.

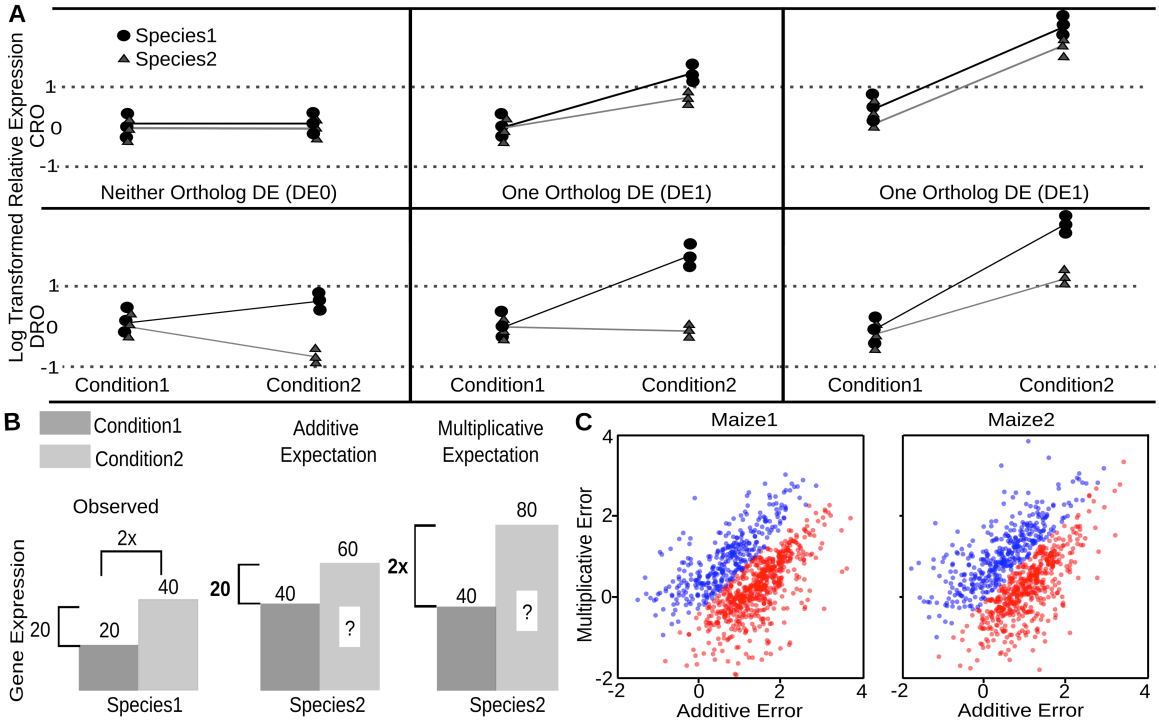

Figure 5. Conceptual illustration of the differentially regulated ortholog model. (A) Illustration of the different classification outcomes that can be produced for a given gene pair using both a DEG-based analysis (testing whether the expression pattern of each gene changes significantly between conditions) and a DRO-based analysis (testing whether the pattern across the two conditions is significantly different between copies of the same gene in both species). **(B)** Two models, additive and multiplicative, for predicting what a conserved pattern of gene regulation should look like when the underlying level of expression changes. **(C)** Relationship between prediction error (log10 transformed) for expression under cold stress using a multiplicative model to predict expression between maize1/maize2 gene pairs or an additive model to predict expression between maize1/maize2 gene pairs. Maize1: Predictions for the expression pattern of maize2 genes using data from their maize1 homeologs; Maize2: Predictions for the expression pattern of maize1 genes using data from their maize2 homeologs. Blue dots mark cases where the additive model was the better predictor; red dots mark cases where the multiplicative model was the better predictor.

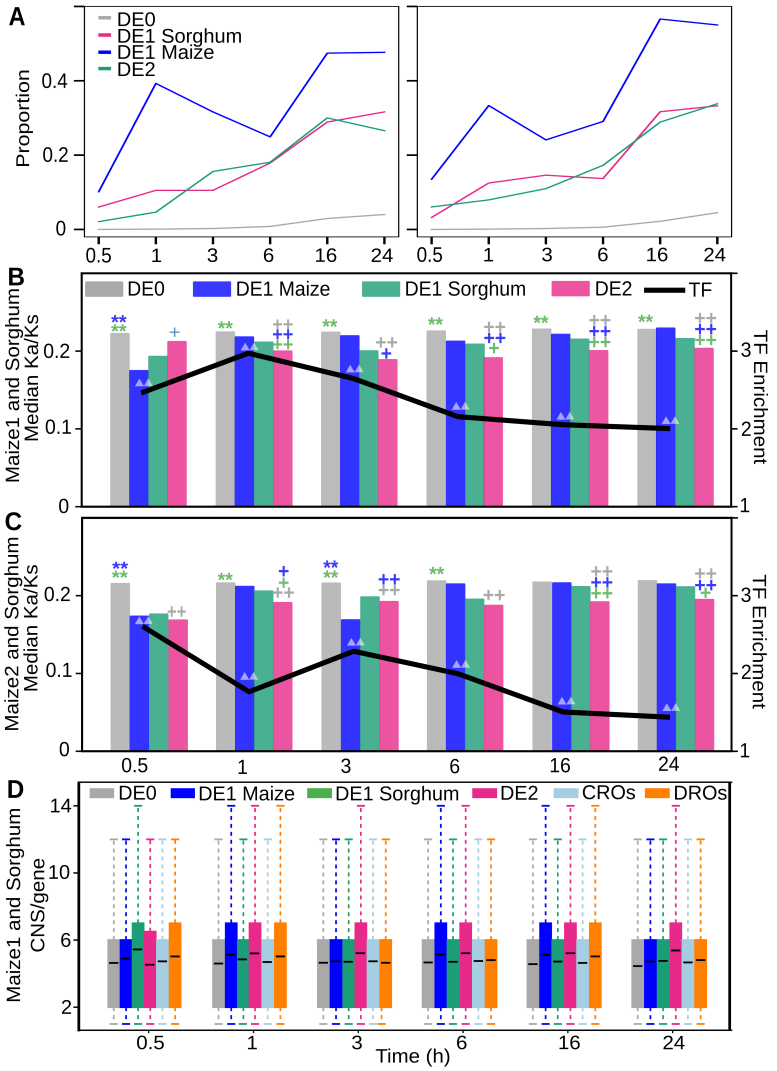

Figure 6. Characteristics of genes in different DEG groups at different timepoints. **(A)** The proportion of gene pairs classified as DROs between maize and sorghum in different DEG groups at each of the six time points examined. **(B)-(C)** Median ratios of non-synonymous substitutions to synonymous substations in coding sequences for maize and sorghum for gene pairs classified as DE0, DE1, or DE2 at each of six time points. Time points where there is a statistically significant difference in Ka/Ks ratio between DE2 and any of the other three categories are marked with either + (if $p<0.05$) or ++ (if $P<0.01$). Color of the + indicates the category to which DE2 is being compared. Time points where there is a statistically significant difference in Ka/Ks ratio between DE0 and either DE1 maize or DE1 sorghum categories are marked with either * (if $p<0.05$) or ** (if $P<0.01$). Color of the * indicates the category to which DE0 is being compared. (B)-(C) Enrichment of genes annotated as transcription factor genes among DE2 gene pairs relative to all syntenic gene pairs indicated by the black line and the right hand axis. Double white triangles mark time points where the enrichment is statistically significant ($p < 0.01$). **(D)** Frequency of CNS within the promoters of genes classified as DE0, DE1 maize, DE1 sorghum, DE2, DRO, or CRO at each of the six time points. Black lines within the box plot mark the average number of CNS per gene for each category. These triangles are difficult to see against the dark background: perhaps you should use white triangles instead.

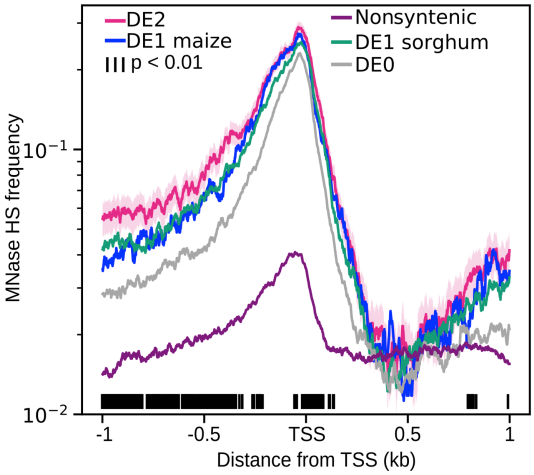

Figure 7. **Chromatin patterns associated with different groups of genes in maize and sorghum.** Patterns of MNase hypersensitive regions around the transcriptional start sites of genes classified based on their pattern of gene regulation in the 24 hour stress time point. Maize1 sorghum gene pairs and maize2 sorghum gene pairs were aggregated to increase statistical power. The lighter band around the DE2 line indicates a two standard deviation confidence interval. Black bars at the bottom of the graph indicate individual base pair positions where the amount of open chromatin associated with DE2 genes is significantly different from that of each of the other four categories displayed with a p-value < 0.01 for each comparison. Pairwise comparisons were performed using a Fisher Exact Test.

Parsed Citations

Anders, S., Pyl, P.T., and Huber, W. (2014). HTSeq-a python framework to work with high-throughput sequencing data. Bioinformatics. 31(2): 166-169.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Yoav, B., Y., Yosef, H., (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Ser B Methodol. 57: 289-300.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Bennetzen, J. L., and Freeling, M. (1993). Grasses as a single genetic system: genome composition, collinearity and compatibility. Trends Genet. 9(8): 259-261.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Burkart-Waco, D., Ngo, K., Lieberman, M., and Comai, L. (2015). Perturbation of parentally biased gene expression during interspecific hybridization. PLoS One. 10(2): e0117293.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Chan, Y.F., Marks, M.E., Jones, F.C., Villarreal, G., Shapiro, M.D., Brady, S.D., Southwick, A.M., Absher, D.M., Grimwood, J., Schmutz, J., Myers, R.M., Petrov D., Jónsson B., Schluter, D., Bell M.A, Kingsley, D.M. (2010). Adaptive evolution of pelvic reduction in sticklebacks by recurrent deletion of a pitx1 enhancer. Science. 327(5963): 302-305.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Chinnusamy, V., Zhu, J., and Zhu, J.K. (2007). Cold stress regulation of gene expression in plants. Trend Plant Sci. 12(10): 444-451.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Chopra, R., Burow, G., Hayes, C., Emendack,Y., Xin, Z., and Burke, J. (2015). Transcriptome profiling and validation of gene based single nucleotide polymorphisms (SNPs) in sorghum genotypes with contrasting responses to cold stress. BMC genomics. 16: 1040.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Chopra, R., Burow, G., Burke, J. J., Gladman, N., and Xin, Z. (2017). Genome-wide association analysis of seedling traits in diverse sorghum germplasm under thermal stress. BMC Plant Biol. 17: 12.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Christin, P.A., Salamin, N., Savolainen, V., Duvall, M.R., and Besnard, G. (2007). C4 photosynthesis evolved in grasses via parallel adaptive genetic changes. Current Biology, 17(14): 1241-1247.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Cooper, G.M., Stone, E.A., Asimenos, G., Green, E.D., Batzoglou, S., and Sidow, A. (2005). Distribution and intensity of constraint in mammalian genomic sequence. Genome Res. 15(7): 901-913.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Cui, S., Huang, F., Wang, J., Ma, X., Cheng, Y., and Liu, J. (2005). A proteomic analysis of cold stress responses in rice seedlings. Proteomics. 5(12): 3162-3172.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Davidson, R. M., Gowda, M., Moghe, G., Lin, H., Vaillancourt, B., Shiu S.-H, Jiang N., and Buell, C.R. (2012). Comparative transcriptomics of three poaceae species reveals patterns of gene expression evolution. Plant J. 71(3): 492-502.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

De Wet J. (1978). Systematics and evolution of sorghum sect. sorghum (gramineae). Am J Bot. 65(4): 477-484.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Eichten, S.R., Swanson-Wagner, R.A., Schnable, J.C, Waters, A.J., Hermanson, P.J., Liu, S., Yeh, CT., Jia, Y., Gendler, K., Freeling, M., Schnable, P.S., Vaughn, M.W., Springer, N.M. (2011). Heritable epigenetic variation among maize inbreds. PLoS Genet. 7(11): e1002372.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Ercoli, L., Mariotti, M., Masoni, A., and Arduini, I. (2004). Growth responses of sorghum plants to chilling temperature and duration of exposure. Eur J Agron. 21(1): 93-103.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Fiedler, K., Bekele, W. A., Matschegewski, C., Snowdon, R., Wieckhorst, S., Zacharias, A., and Uptmoor, R. (2016). Cold tolerance during juvenile development in sorghum: a comparative analysis by genomewide association and linkage mapping. Plant Breeding., 135(5): 598-606.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Freeling, M., Rapaka, L., Lyons, E., Pedersen, B., and Thomas, B.C. (2007). G-boxes, bigfoot genes, and environmental response: characterization of intragenomic conserved noncoding sequences in arabidopsis. Plant Cell. 19(5): 1441-1457.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Freeling, M., Woodhouse, M.R., Subramaniam, S., Turco, G., Lisch, D., and Schnable, J.C. (2012). Fractionation mutagenesis and similar consequences of mechanisms removing dispensable or less-expressed dna in plants. Curr Opin Plant Biol. 15(2): 131-139.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Haibao, T., Klopfenstein, D., Pedersen, B., Flick, P., Sato, K., Ramirez, F., Yunes, J., and Mungall, C. (2015). Goatools: Tools for gene ontology. Zenodo.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Harris R.S. (2007). Improved pairwise alignment of genomic DNA. Ph.D. thesis Pennsylvania State University.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Hetherington, S. E., He, J., and Smillie, R.M. (1989). Photoinhibition at low temperature in chilling-sensitive and-resistant plants. Plant Physiol. 90(4): 1609-1615.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Hollister, J.D. and Gaut, B.S. (2009). Epigenetic silencing of transposable elements: a trade-off between reduced transposition and deleterious effects on neighboring gene expression. Genome Res. 19(8): 1419-1428.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Hollister J.D., Smith L.M., Guo, Y.-L., Ott, F., Weigel, D., and Gaut, B.S. (2011). Transposable elements and small RNAs contribute to gene expression divergence between Arabidopsis thaliana and Arabidopsis lyrata. Proc. Natl. Acad. Sci. U.S.A. 108(6): 2322-2327.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Lai, X., Behera, S., Liang, Z., Lu, Y., Deogun, J. S., and Schnable, J. C. (2017). Stag-CNS: An order-aware conserved non-coding sequences discovery tool for arbitrary numbers of species. Mol Plant.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Lassmann, T., and Sonnhammer, E.L. (2005). Kalign-an accurate and fast multiple sequence alignment algorithm. BMC bioinformatics. 6(1): 298.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Li, L., Briskine, R., Schaefer, R., Schnable, P.S., Myers, C.L., Flagel, L.E., Springer, N.M., and Muehlbauer, G.J. (2016). Co-expression network analysis of duplicate genes in maize (Zea mays L.) reveals no subgenome bias. BMC genomics, 17(1): 875.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Liang, Z. and Schnable, J.C. (2016). RNA-Seq based analysis of population structure within the maize inbred B73. PloS One. 11(6):

e0157942.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-Seq data with DESeq2. Genome Biol. 15(12): 550.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Lovell, J. T., Schwartz, S., Lowry, D. B., Shakirov, E. V., Bonnette, J. E., Weng, X., Wang, M., Johnson, J., Sreedasyam, A., Plott, C., et al. (2016). Drought responsive gene expression regulatory divergence between upland and lowland ecotypes of a perennial c4 grass. Genome Res. 26(4): 510-518.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Makarevitch, I., Waters, A.J., West, P.T., Stitzer, M., Hirsch, C.N., Ross-Ibarra, J., and Springer, N.M (2015). Transposable elements contribute to activation of maize genes in response to abiotic stress. PLoS Genet. 11(1): e1004915.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal. 17(1):10-12.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Muiño, J.M., de Bruijn, S., Pajoro, A., Geuten, K., Vingron, M., Angenent, G.C., and Kaufmann, K. (2016). Evolution of dna-binding sites of a floral master regulatory transcription factor. Mol Biol Evol. 33(1): 185-200.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Naito, K., Zhang, F., Tsukiyama, T., Saito, H., Hancock, C.N., Richardson, A.O., Okumoto, Y., Tanisaka, T., and Wessler, S.R. (2009). Unexpected consequences of a sudden and massive transposon amplification on rice gene expression. Nature. 461(7267): 1130-1134.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Ng, P.C and Henikoff, S. (2001). Predicting deleterious amino acid substitutions. Genome Res. 11(5): 863-874.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Olsen, J., McMahon, C., and Hammer, G. (1993). Prediction of sweet corn phenology in subtropical environments. Agron J. 85(2): 410-415.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Orr, H.A. (1998). The population genetics of adaptation: the distribution of factors fixed during adaptive evolution. Evolution. 52(4): 935-949.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Orr, H.A., (1999). The evolutionary genetics of adaptation: a simulation study. Genet Res. 74(03): 207-214.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Paschold, A., Larson, N.B., Marcon, C., Schnable, J.C., Yeh C.-T., Lanz, C., Nettleton, D., Piepho, H.-P., Schnable, P.S., and Hochholdinger, F. (2014). Nonsyntenic genes drive highly dynamic complementation of gene expression in maize hybrids. Plant Cell. 26(10): 3939-3948.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Paterson, A.H., Bowers J.E., Bruggmann, R., Dubchak, I., Grimwood, J., Gundlach, H., Haberer, G., Hellsten, U., Mitros, T., Poliakov, A., et al. (2009). The Sorghum bicolor genome and the diversification of grasses. Nature. 457(7229): 551-556.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Pearce, R. S., (2001). Plant freezing and damage. Annals of Botany, 87(4): 417-424.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Priest, H.D., Fox, S.E., Rowley, E.R., Murray, J.R., Michael, T.P., and Mockler, T.C. (2014). Analysis of global gene expression in brachypodium distachyon reveals extensive network plasticity in response to abiotic stress. PLoS One. 9(1): e87499.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Quinn, P. J., (1988). Effects of temperature on cell membranes. Symp Soc Exp Biol. 42: 237-258.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Reva, B., Antipin, Y., and Sander, C. (2011). Predicting the functional impact of protein mutations: application to cancer genomics. Nucleic Acids Res. 39(17): e118.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Revilla, P., Rodríguez, V.M., Ordás, A, Rincent, R., Charcosset, A., Giauffret, C., Melchinger, A.E., Schön, C.-C., Bauer, E., Altmann, T., et al. (2016). Association mapping for cold tolerance in two large maize inbred panels. BMC Plant Biol. 16: 127.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W., and Smyth G.K. (2015). Limma powers differential expression analyses for RNA-Sequencing and microarray studies. Nucleic Acids Res. 43 (7): e47
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Robinson M. D, McCarthy D. J, and Smyth G. K 2010. edger: a bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics, 26(1): 139-140.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Rodgers-Melnick, E., Vera D.L., Bass, H.W., and Buckler, E.S. (2016). Open chromatin reveals the functional maize genome. Proc. Natl. Acad. Sci. U.S.A. 113(22): E3177-E3184.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Sánchez, B., Rasmussen, A., and Porter, J.R. (2014). Temperatures and the growth and development of maize and rice: a review. Global Change Biol. 20(2): 408-417.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Schaefer, R.J., Briskine, R., Springer, N.M., and Myers, C.L. (2014). Discovering functional modules across diverse maize transcriptomes using cob, the co-expression browser. PloS one, 9(6): e99193.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Schnable, J. C., (2015). Genome evolution in maize: From genomes back to genes. Annu Rev Plant Biol. 66: 329-343.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Schnable, J.C, Springer, N.M., and Freeling, M. (2011). Differentiation of the maize subgenomes by genome dominance and both ancient and ongoing gene loss. Proc Natl Acad Sci U.S.A. 108(10): 4069-4074.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Schnable J.C., Freeling, M., and Lyons, E. (2012). Genome-wide analysis of syntenic gene deletion in the grasses. Genome Biol Evol. 4(3): 265-277.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Schnable, P.S., Ware, D., Fulton, R.S., Stein, J.C., Wei, F., Pasternak, S., Liang, C., Zhang, J., Fulton, L., Graves, T.A, et al. 2009. The B73 maize genome: complexity, diversity, and dynamics. Science. 326(5956): 1112-1115.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Schurch, N. J., Schofield, P., Gierlinski, M., Cole, C., Sherstnev, A., Singh, V., Wrobel, N., Gharbi, K., Simpson, G. G., Owen-Hughes, T., et al. (2016). How many biological replicates are needed in an RNA-Seq experiment and which differential expression tool should you use? RNA. 22(6): 839-851.

Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Shantz, H. (1954). The place of grasslands in the earth's cover. Ecology. 35(2): 143-145.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Shaykewich, C. (1995). An appraisal of cereal crop phenology modelling. Can J Plant Sci. 75(2): 329-341.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Studer, A., Zhao, Q., Ross-Ibarra, J., and Doebley, J. (2011). Identification of a functional transposon insertion in the maize domestication gene tb1. Nature Genet. 43(11): 1160-1163.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Swigonová, Z., Lai, J., Ma, J., Ramakrishna, W., Llaca, V., Bennetzen, J.L., and Messing, J. (2004). Close split of sorghum and maize genome progenitors. Genome Res. 14(10a): 1916-1923.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Tenhaken, R., (2014). Cell wall remodeling under abiotic stress. Front. Plant Sci.,
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

5. doi: 10.3389/fpls.2014.00771.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Tiwari, S., Krishnamurthy, S., Kumar, V., Singh, B., Rao, A., SV, A.M., Rai,V., Singh, A.K., and Singh, N.K. (2016). Mapping qtls for salt tolerance in rice (Oryza sativa L.) by bulked segregant analysis of recombinant inbred lines using 50k snp chip. PloS One., 11(4): e0153610.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Torres-Oliva, M., Almudi, I., McGregor, A.P., and Posnien, N. (2016). A robust (re-) annotation approach to generate unbiased mapping references for RNA-Seq-based analyses of differential expression across closely related species. BMC genomics, 17(1): 392.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Turco, G., Schnable, J.C., Pedersen, B., and Freeling, M. (2013). Automated conserved non-coding sequence (CNS) discovery reveals differences in gene content and promoter evolution among grasses. Front Plant Sci. 4: 170.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Usadel, B., Poree, F., Nagel, A., Lohse, M., CZEDIK-EYSENBERG A., and Stitt, M. (2009). A guide to using mapman to visualize and compare omics data in plants: a case study in the crop species, maize. Plant Cell Environ. 32(9): 1211-1229.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

van Heerwaarden, J., Doebley, J., Briggs, W.H., Glaubitz, J.C., Goodman, M.M., Gonzalez, J.d.J.S., and Ross-Ibarra, J. (2011). Genetic signals of origin, spread, and introgression in a large sample of maize landraces. Proc Natl Acad Sci U.S.A. 108(3): 1088-1092.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Vera, D.L., Madzima, T.F., Labonne, J.D., Alam M.P., Hoffman, G.G., Girimurugan, S., Zhang, J., McGinnis, K.M., Dennis, J.H., and Bass, H.W. (2014). Differential nuclease sensitivity profiling of chromatin reveals biochemical footprints coupled to gene expression and functional dna elements in maize. Plant Cell. 26(10): 3883-3893.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Wallace, J., Larsson, S., and Buckler, E. (2014). Entering the second century of maize quantitative genetics. Heredity, 112(1): 30-38.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Wang, L., Czedik-Eysenberg, A., Mertz, R. A., Si, Y., Tohge, T., Nunes-Nesi, A., Arrivault, S., Dedow, L. K., Bryant, D. W., Zhou, W., et al. (2014). Comparative analyses of C4 and C3 photosynthesis in developing leaves of maize and rice. Nature Biotechnol. 32(11): 1158-1165.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Waters, A.J, Makarevitch, I., Noshay, J., Burghardt, L.T, Hirsch, C.N, Hirsch, C.D., and Springer, N.M. (2017). Natural variation for gene expression responses to abiotic stress in maize. Plant J. 89: 706-717.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Wendorf, F., Close, A.E., Schild, R., Wasylikowa, K., Housley, R.A., Harlan, J.R., and Krolik, H. (1992). Saharan exploitation of plants 8,000 years bp. Nature, 359(6397): 721-724.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Wu, T.D and Nacu, S. (2010). Fast and snp-tolerant detection of complex variants and splicing in short reads. Bioinformatics. 26(7): 873-881.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Wu, T.D and Watanabe, C.K (2005). Gmap: a genomic mapping and alignment program for mRNA and EST sequences. Bioinformatics. 21(9): 1859-1875.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Xu, G., Wang, X., Huang, C., Xu, D., Li, D., Tian, J., Chen, Q., Wang, C., Liang, Y., Wu, Y., et al. (2017). Complex genetic architecture underlies maize tassel domestication. New Phytol. 214: 852-864.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Yang, Z. (2007). Paml 4: phylogenetic analysis by maximum likelihood. Mol. Biol. Evol. 24(8): 1586-1591.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Yilmaz, A., Nishiyama, M.Y., Fuentes, B.G., Souza, G.M., Janies, D., Gray, J., and Grotewold, E. (2009). Grassius: a platform for comparative regulatory genomics across the grasses. Plant Physiol. 149(1): 171-180.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Zhang, W., Wu, Y., Schnable, J. C., Zeng, Z., Freeling, M., Crawford, G. E., and Jiang, J. (2012). High-resolution mapping of open chromatin in the rice genome. Genome Res. 22(1): 151-162.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Zhang, Y., Ding, Z., Ma, F., Chauhan, R.D., Allen, D.K., Brutnell, T.P., Wang, W., Peng, M., and Li, P. (2015). Transcriptional response to petiole heat girdling in cassava. Sci Rep. 5: 8414.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title